

Safety Quantification for Nonlinear and Time-Delay Systems using Occupation Measures

A Dissertation Presented
by

Jared Franklin Miller

to

The Department of Electrical and Computer Engineering

in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Electrical Engineering

**Northeastern University
Boston, Massachusetts**

April 2023

To my parents, Wayne and Debbie.

Acknowledgments

I would like to begin by thanking my parents, Debbie Mitzner and Wayne Miller, for all of their advice and support. Wayne Miller voluntarily acted as a copy-editor for all of my publications, including this thesis.

My next thanks go to my advisor, Mario Sznaier, at the Robust Systems Laboratory in Northeastern. I met Mario in Fall of 2016 as a BS/MS student, when I asked him for an override in order to attend his Spring 2017 special topics course in Big Data, Sparsity, and Control. During the opening lecture, he gave a survey of sparse methods in robust regression, computer vision, and system identification. At the close of this first lecture, I asked him how I could learn more and join his lab, and he responded by giving me the door code. I officially joined the lab as a PhD student in Fall of 2018, and then acted as a teaching assistant for his Sparsity course in 2019, 2021, and 2022. Mario introduced me to topics such as rank-minimization, sparsity, Frank-Wolfe methods, measures, polynomial optimization, and data-driven control. Mario's wealth of experience allows him to draw connections between different disciplines and rapidly determine if ideas have promise, have been done before, and/or are marginal/intractable. His discussions with myself and other members of the laboratory help us refine our approaches and experiments, providing guidance in sometimes unexpected ways, and leading us into new directions of research. I would like to thank Octavia Camps (co-director of the Robust Systems Laboratory) for the discussions and application of my methods towards computer vision.

I would like to thank all of my colleagues in the Robust Systems Lab over the course of my ~6 years there. Specific recognition goes to my collaborators Tianyu Dai, Rajiv Singh, Jian Zheng, Armand Comas, Biel Roig-Solvas, and Zachary Walker-Liang.

There are many other students and professors at Northeastern (outside the Robust Systems Laboratory) to thank. Students in research include Muhammad Ali Al-Radhawi, Aria Masoomi, Tooba Imtiaz, Arthur Costello Branco de Oliveira, Chieh Wu, Zulqarnain Khan, and James Massucco.

Professors at Northeastern to thank include Bahram Shafai, Eduardo Sontag, Milad Siami, Jennifer Dy, Dave Rosen, David Kaeli, and Felipe Pait. Bahram Shafai was my Capstone project advisor in 2017-2018, and was the coauthor for my first conference paper (about the vertical motion of air cushion vehicles) [1]. Eduardo Sontag taught a Special Topics course about synthetic biology in Spring 2018, during which my final project analyzed orthogonal ribosome models for protein translation. I worked with Eduardo and Muhammad on this project in the time between my MS graduation and my PhD beginning (publishing our findings at [2]), and my discussions with Eduardo's research group helped me understand how to apply systems theory methods towards problems in biology and chemistry.

When I was deciding between undergraduate institutions in 2013, David Kaeli personally called me to recruit me to Northeastern. His Embedded Systems pilot course was the first time I watched control algorithms that I wrote being deployed onto robotic systems (FPGA via Simulink or C).

I would like to thank Yang Zheng and Antonis Papachristodoulou for teaching me how to formulate research questions and to write papers. I visited Oxford in July 2018 before my PhD, during which I shared an idea about structured rank-minimization under chordal sparsity. This turned into my first conference publication in the theory of control [3]. When in February of 2019 I realized that positive semidefinite approximations could be combined with structure, I again worked with Yang, Mario, and Antonis to develop these findings for efficient convex optimization. This work in Decomposed Structured Subsets became my first journal article [4]. Yang was extremely helpful with respect to organizing my research ideas into theorems and statements, which guided me through my work and the layout/presentation of this thesis.

The first time I presented my control research at an in-person conference was at CDC 2019 in Nice, France. Following this presentation, Didier Henrion and Victor Magron from LAAS-CNRS asked me questions about my work in rank-minimization [3] (after a suggestion to attend from Mario). Once I returned to the United States, I received an email from Didier inviting me to apply for a Chateaubriand Fellowship. I was awarded this fellowship with a scheduled time-frame of January-July, 2021, but we delayed this multiple times due to the ongoing COVID epidemic. During these delays, I worked remotely with Didier Henrion, Milan Korda, and Victor Magron in a ‘virtual’ fellowship, and presented my work twice at BrainPOP seminars. I stayed physically at LAAS between January-July, 2022, during which I met and fruitfully worked with the POP and MAC teams of LAAS-CNRS. I also got to interact with many PhDs at LAAS-CNRS during my 2022 visit, with special thanks to Corbinian Schlosser, Vit Cibulka, Alexey Lazayev, Nicola Zaupa, Antonio

Bellin, Loi Do, Olga Yufuvera, Filip Becanovic, and Anh Le. Some of the highlights of my time at LAAS-CNRS included the weekly Sunday visits to Marché St. Aubin (a wonderful street market).

My research is the product of interactions and discussions with several research groups. Specific thanks goes to Fred Leve (Air Force Office of Scientific Research), Yang Zheng (UC San Diego), Antonis Papachristodoulou (University of Oxford), Ufuk Topcu (UT Austin), Necmiye Ozay (University of Michigan), Matteo Tacchi (GIPSA), Ashkan Jasour (Caltech/JPL), S. Farokh Atashzar (NYU), Colin Jones (EPFL), Simone Naldi (University of Limoges), Mauricio Velasco (Universidad de los Andes), Jesús A. de Loera (UC Davis), Daniel Brosch (Alpen-Adria-Universität Klagenfurt), Alain Rapaport (Marseille), Chiara Meroni (ICERM), and Kwesi Rutledge (MIT).

My penultimate thanks goes to Roy Smith, John Lygeros, and Florian Dörfler at the Automatic Control Laboratory (IfA) in ETH, Zurich. I look forward to working with the members of IfA, maintaining my current connections, and fostering new collaborations as part of my next steps.

I would like to close these acknowledgements by thanking the Warden (Figure 0.1). When I lived in Toulouse between January and July of 2022, I would sometimes be woken up by a repetitive banging noise against the windowpane. This is how I met the Warden, a pigeon (Rock Dove) that had a nest on the windowsill outside my apartment. I tested positive for COVID-19 on March 29, 2022, and the Warden kept watch to make sure I stayed in my apartment the entire 7 days of quarantine.



Figure 0.1: The Warden

Thank you again to all of the people who have helped me over the course of my Control career.

Contents

Acknowledgments	ii
List of Figures	viii
List of Tables	xi
List of Acronyms	xii
Abstract of the Dissertation	xiv
1 Introduction	1
1.1 Motivation	1
1.2 Summary	3
1.3 Publications	4
2 Preliminaries	6
2.1 Notation	6
2.2 Measure Theory	6
2.3 Occupation Measures	9
2.4 Moment-SOS Hierarchy	11
Part 1: Peak Estimation and Safety Analysis	15
3 Peak Estimation and Recovery	16
3.1 Introduction	16
3.2 Peak Estimation Programs	17
3.3 Recovery	19
3.4 Recovery Examples	21
3.5 Global Attractor Peak Recovery	23
3.6 Conclusion	28
4 Peak Estimation under Uncertainty	29
4.1 Introduction	29
4.2 Uncertainty Models	30
4.3 Continuous-Time Uncertain Peak Estimation	32

4.4	Discrete-Time Uncertain Peak Estimation	39
4.5	Conclusion	44
5	Peak Estimation for Safety Analysis	45
5.1	Introduction	45
5.2	Barrier Functions	47
5.3	Safety Margins and Maximin Peak Estimation	48
5.4	Distance Estimation Program	52
5.5	Finite-Dimensional Programs	57
5.6	Exploiting Correlative Sparsity	61
5.7	Shape Safety	64
5.8	Numerical Examples	69
5.9	Extensions	73
5.10	Conclusion	76
Part 2: Robust Safety and Peak Minimizing Control		78
6	Robust Counterparts and Data Driven Analysis	79
6.1	Background	79
6.2	Preliminaries	81
6.3	Analysis and Control Problems	84
6.4	Decomposed Lie Constraint	89
6.5	Polynomial Approximation	90
6.6	Data-Driven Setting	94
6.7	Examples	96
6.8	Conclusion	102
7	Safety Quantification using Peak Minimizing Control	106
7.1	Introduction	106
7.2	Peak Minimizing Control	108
7.3	Crash-Safety Program	108
7.4	Robust Crash-Safety and Data-Driven Analysis	112
7.5	SOS Programs	114
7.6	Subvalue Map	116
7.7	Examples	119
7.8	Conclusion	123
8	Distance-Maximizing Control	126
8.1	Distance Maximizing Control Program	128
8.2	Robust Formulation for Distance-Maximizing Control	132
8.3	Shape-Distance Maximization	135
8.4	Conclusion	137
Part 3: Peak Estimation Extensions to non-ODE Systems		139

9 Peak Estimation for Hybrid Systems	140
9.1 Introduction	140
9.2 Hybrid Systems Preliminaries	141
9.3 Peak estimation of hybrid systems	143
9.4 Extensions	150
9.5 Numerical Examples	151
9.6 Conclusion	155
10 Value-at-Risk Peak Estimation	158
10.1 Introduction	158
10.2 Preliminaries	159
10.3 Peak Value-at-Risk Estimation for Stochastic Systems	162
10.4 Finite Moment Program	167
10.5 Extensions	169
10.6 Numerical Examples	172
10.7 Conclusion	177
11 Peak Estimation for Time-Delay Systems	178
11.1 Introduction	178
11.2 Time Delay Background	180
11.3 Peak Linear Programs	182
11.4 Peak Moment Program	187
11.5 Extensions	190
11.6 Numerical Examples	195
11.7 Conclusion	201
12 Conclusion	202
Bibliography	204
A Appendices	222
A.1 Proof of Strong Duality for Distance Estimation in Theorem 5.4.5	222
A.2 Continuity of Multipliers	223
A.3 Integral Costs and Robust Counterparts	227
A.4 Polynomial Approximation of the Auxiliary Function	230
A.5 Polynomial Approximation of Multipliers	230
A.6 Robust Duality and Recovery	233
A.7 Strong Duality of Linked Semidefinite-Measure Programs	238
A.8 Strong Duality of Chance-Peak Linear Programs	242
A.9 Delay Structures	243
A.10 Proof of Strong Duality for DDE Peak Estimation	247
A.11 Subvalue Functionals and DDE Control	249
A.12 Improved Accuracy of DDE Control Problems	256
A.13 Joint + Component Measure	260

List of Figures

0.1	The Warden	iv
1.1	Safety of the car is quantified by a 5" distance of closest approach to the tree	1
1.2	Credit to Mayo Clinic News Network	2
2.1	Dirac Delta: solid boxes contain x' , dotted boxes miss x'	7
2.2	Trajectories of $x' = -x(x + 2)(x - 1)$ and boxes	11
3.1	Minimizing x_2 along Flow system (3.9)	22
3.2	Maximize $\ x\ _2^2$ along (3.10)	23
3.3	Maximize $h = 1 - \cos \theta$ along pendulum with friction	24
3.4	Maximize x_1^2 on Van-der-Pol system in (3.20)	27
3.5	Maximize $25\tilde{x}_1$ on the scaled Lorenz system in (3.21)	28
4.1	Maximize x_1 at order 4 with $w(t) \in W$	37
4.2	Maximize x_2 on three-wave system (4.13)	38
4.3	Maximum angle for 1DOF attitude controller	39
4.4	Minimize x_2 on system (4.24)	44
5.1	Trajectories of Flow system (3.9)	46
5.2	Degree-6 Barrier function for Flow system (3.9)	47
5.3	Peak analysis of system (5.7) at $d = 3$	50
5.4	The value of $\min(x_1, x_2)$ along trajectories (5.7)	50
5.5	Flow system is safe, $p^* \leq -0.2831$	51
5.6	Safety margin scaling contours	51
5.7	Flow system L_2 bound of 0.2831 with optimal trajectory recovery	53
5.8	Comparison between L_2 distance and safety margin contours	53
5.9	Correlative Sparsity Pattern (CSP) with 4-States and Chordal Extension	63
5.10	Shape moving and rotating along Flow (3.9) trajectories	65
5.11	Collision if X_u is relaxed to its convex hull.	70
5.12	Moon unsafe-set has an L_2 bound of 0.1592	70
5.13	Distance contours at order-5 relaxation for the Twist system (5.34)	71
5.14	Translation, L_2 bound of 0.1465	73
5.15	Rotation, L_2 bound of 0.1425	74
5.16	Uncertain Flow (5.40), L_2 bound of 0.1691	75

5.17 Flow system L_1 bound of 0.4003	76
6.1 Plot of Uncertain Flow system (6.15) trajectories	86
6.2 Observed and Corrupted Susceptible, Infected, Removed (SIR)	97
6.3 Active and inactive constraints describing Θ_D in Figure 6.2b	98
6.4 Maximum value of $I(t)$ over a time horizon of $T = 40$, with unknown (β, γ)	98
6.5 Order-5 bound on minimal x_2 for Flow (6.46) under ellipope-constrained noise	99
6.6 Observed data of Flow system (3.9) within a circle	100
6.7 Minimizing x_2 on Flow system (3.9) at order-4 SOS tightening	101
6.8 Distance estimate of (3.9) at order 5	102
6.9 100 observations of Twist system (5.34)	103
6.10 Twist (5.34) system where either B^1 or B^3 are unknown.	104
6.11 Twist (5.34) with unknown (B^1, B^3)	104
6.12 Reachable set estimation of twist (5.34) system where B^1 is known and B^3 are unknown.	105
6.13 Order 6 Region of Attraction (ROA) for controlled Flow (3.9)	105
7.1 Two trajectories with nearly the same distance but different crash-bounds	109
7.2 Subvalue function for Flow system (7.4) between degrees 1..5.	120
7.3 Numerical optimal control yields worst-case $Q^* \approx 0.4639$ for the half-circle X_u . .	121
7.4 Numerical optimal control yields $Q^* \approx 0.3232$ for the moon X_u	122
7.5 Subvalue map for the moon (7.27) on the flow system (7.4)	123
7.6 Subvalue for data-driven (3.9) between degrees 1..4	124
7.7 Numerically computed crash-bound for data-driven Flow (3.9)	125
7.8 Applied control for the data-driven Flow crash system.	125
8.1 Distance maximizing control avoiding a rectangular block	127
9.1 Deterministic Two-Mode Bound of $x_1^2 \leq 0.3958 = p_5^*$	152
9.2 Noisy Two-Mode Bound of $x_1^2 \leq 0.9660 = p_5^*$	153
9.3 Deterministic Two-Mode Distance Bound of $\min_{y \in X_u} \ x - y\ _2 \leq 0.0891 = \sqrt{p_5^*}$	154
9.4 Peak Estimation of Right-Left Wrap Dynamics (9.20) and (9.21)	155
9.5 Distance bound of $\inf_{y \in X_u} \ x - y\ _2 \leq 0.2877$	156
10.1 Trajectories of (10.30) with $\epsilon = \{0.5, 0.15\}$ bounds	173
10.2 Trajectories of (10.31) with $\epsilon = 0.5$ (transparent red) and $\epsilon = 0.15$ (solid red) bounds	174
10.3 Trajectories of (10.32) with $\epsilon = \{0.5, 0.15\}$ bounds	175
10.4 Trajectories of (10.33) with $\epsilon = \{0.5, 0.15\}$ bounds	176
11.1 Continuity of (11.2) trajectories increases every $\tau = 1$ time step	180
11.2 All histories pass through $x_h(0) = 1$	181
11.3 Bifurcation of stability as τ exceeds $\pi/2$ in $\dot{x} = -x(t - \tau)$	181
11.4 The same μ_h leads to different trajectories in times $(0, 5]$	186
11.5 Constant histories in the black box	192
11.6 Peak infected population vs. time delay	196
11.7 SIR peak estimation and recovery at order 3	197

11.8 Comparison of delayed Flow systems (11.30) with lags $\tau = 0$ and $\tau = 0.75$ in times $t \in [0, 20]$	198
11.9 Minimize x_2 on the delayed Flow system (11.30)	199
11.10 Minimize $c(x; X_u)$ on the delayed Flow system (11.30)	199
11.11 Maximize x_1 on the delayed time-varying (11.31)	200
11.12 A 3d plot of (11.31) and its x_1 bound	200
A.1 Open and closed-loop trajectory with control	256
A.2 Auxiliary function v and value functional V from (A.102)	256

List of Tables

5.1	Characterization of optimal trajectory in distance estimation	52
5.2	Sizes of moment matrices in Linear Matrix Inequality (LMI) (5.18)	59
5.3	L_2 bounds for the Twist Example	72
5.4	Time in seconds for the Twist Example	72
6.1	Size of largest Lie constraint Gram Matrix (Peak Estimation)	81
7.1	Crash-bounds at $X_0 = [1; 0]$ under Sum of Squares (SOS) tightenings	120
7.2	Crash-bounds at $X_0 = [0; 0]$ for the moon (7.27) under SOS tightenings	121
7.3	Data-Driven Crash-bounds at $X_0 = [1; 0]$ under SOS tightenings	123
9.1	Sizes of moment matrices in LMI (9.13)	149
10.1	Size of Moment Matrices in LMI (10.23)	169
10.2	Chance-Peak estimation of the Stochastic Flow System (10.30) to maximize $p(x) = -x_2$	172
10.3	Solver time (seconds) to compute Table 10.2	172
10.4	Chance-Peak estimation of the Stochastic Twist System (10.31) to maximize $p(x) = x_3$	173
10.5	Solver time (seconds) to compute Table 10.4	174
10.6	Chance-Peak Exit-Statistic estimation of Standard Brownian Motion System . . .	174
10.7	Chance-Peak squared distance lower bounds for System (10.32)	175
10.8	Solver time (seconds) to compute Table 10.7	176
10.9	Chance-Peak upper-bounds for $p(x) = -x_2$ for the Switched System (10.33) . .	177
10.10	Solver time (seconds) to compute Table 10.9	177
11.1	Size of Moment Matrices in LMI (11.15)	190
A.1	Semidefinite Program (SDP) approximation bounds to program (A.94)	255
A.2	Comparison of (11.11) and (A.127) SDP bounds for the delayed Flow system . . .	262
A.3	Time (seconds) to obtain SDP bounds in Table A.2	263

List of Acronyms

BSA	Basic Semialgebraic
BRS	Backward Reachable Set
DDE	Delay Differential Equation
CSP	Correlative Sparsity Pattern
HJB	Hamilton-Jacobi-Bellman
LMI	Linear Matrix Inequality
LP	Linear Program
MV	Measure-Valued
OCP	Optimal Control Problem
ODE	Ordinary Differential Equation
PC	Piecewise Continuous
PDE	Partial Differential Equation
PMI	Polynomial Matrix Inequality
PSD	Positive Semidefinite
PD	Positive Definite
ROA	Region of Attraction
SDE	Stochastic Differential Equation
SDP	Semidefinite Program
SDR	Semidefinite Representable
SIR	Susceptible, Infected, Removed
SOC	Second-Order Cone

SOCP Second-Order Cone Program

SOS Sum of Squares

TV Total Variation

VaR Value-at-Risk

VP Vysochanskij-Petunin

WSOS Weighted Sum of Squares

Abstract of the Dissertation

Safety Quantification for Nonlinear and Time-Delay Systems
using Occupation Measures

by

Jared Franklin Miller

Doctor of Philosophy in Electrical Engineering (CCSP)

Northeastern University, April 2023

Dr. Mario Sznaier, Advisor

This research extends an occupation measure framework to analyze and quantify the safety of dynamical systems. A motivating application of trajectory analysis is in peak estimation, which finds the extreme values of a state function along trajectories. Examples of peak estimation include finding the maximum height of a wave, voltage on a power line, speed of a vehicle, and infected population in an epidemic. Peak estimation can be applied towards safety quantification, such as by measuring the safety of a trajectory by its distance of closest approach to an unsafe set.

A finite-dimensional but nonconvex peak estimation problem can be converted into an infinite-dimensional linear program (LP) in measures, which is in turn bounded by a convergent sequence of semidefinite programs. The LP is posed in terms of an initial, a terminal, and an occupational measure, where the occupation measure contains all possible information about the dynamical systems' trajectories. This research applies measure-based methods towards safety quantification (e.g. distance estimation, control effort needed to crash), hybrid systems, bounded-uncertain systems (including for data-driven analysis), stochastic systems, and time-delay systems. The modularity of this measure-based framework allows for multiple problem variations to be applied simultaneously (e.g., distance estimation under time-delays), and for optimization models to be synthesized using MATLAB. Solving these optimization problems results in certifiable guarantees on system performance and behavior.

Chapter 1

Introduction

1.1 Motivation

The motivation for this thesis is to provide methods for safety quantification and trajectory analysis. Figure 1.1 is a visual example of the importance of quantifying the safety of trajectories.

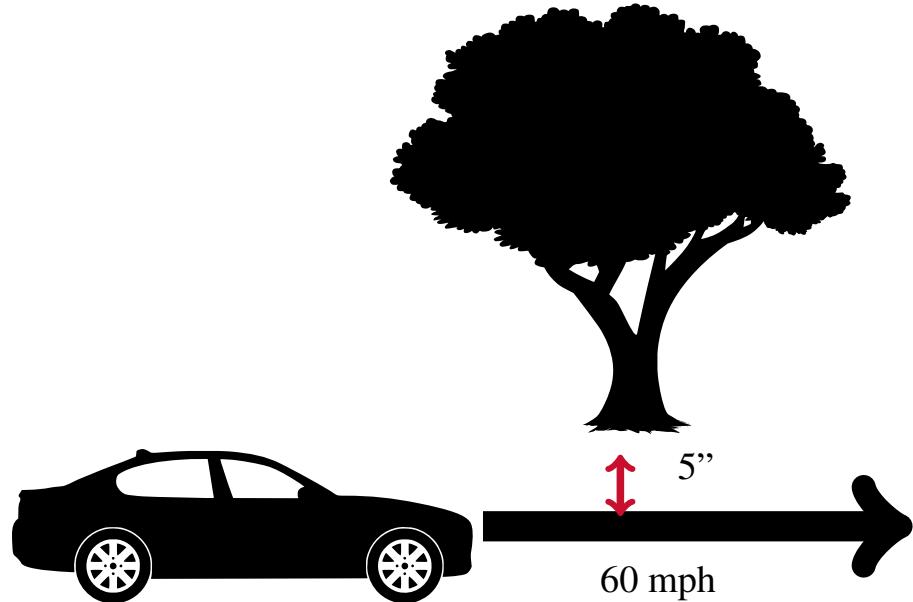


Figure 1.1: Safety of the car is quantified by a 5" distance of closest approach to the tree

An agent is driving a car at a speed of 60 mph. In the path of the car there is a tree (unsafe set/obstacle), and the agent would like to evaluate whether their current course is appropriate. If

CHAPTER 1. INTRODUCTION

the agent receives notification that the current driving plan is safe (the car would not crash into the tree), then the agent will continue its control scheme. If the agent is provided the additional information that the car will attain a 5" clearance to the tree when traveling at 60 mph, the agent would likely deem their current control scheme unsafe and would find a new plan. The distance of closest approach of 5" quantified the safety of the current (verified-safe) controller and allowed for assessment of the control model.

Safety is quantified in Figure 1.1 by finding (or lower-bounding) the distance of closest approach. This distance estimation problem is a particular instance of a peak estimation problem [5, 6]. Using techniques from optimal control theory, the peak estimation problem can be approximated by a convergent sequence of Semidefinite Programs (SDPs) based on occupation measure Linear Programs (LPs) [7, 8]. This thesis extends the peak estimation framework of occupation measures towards other methods of systems analysis with differing dynamical behavior.

This thesis originated during the COVID-19 pandemic in March-April 2020. The phrase ‘flatten the curve’ had entered the public lexicon, in which the peak infection rate must not exceed the hospital system’s capacity for care (Figure 1.2).

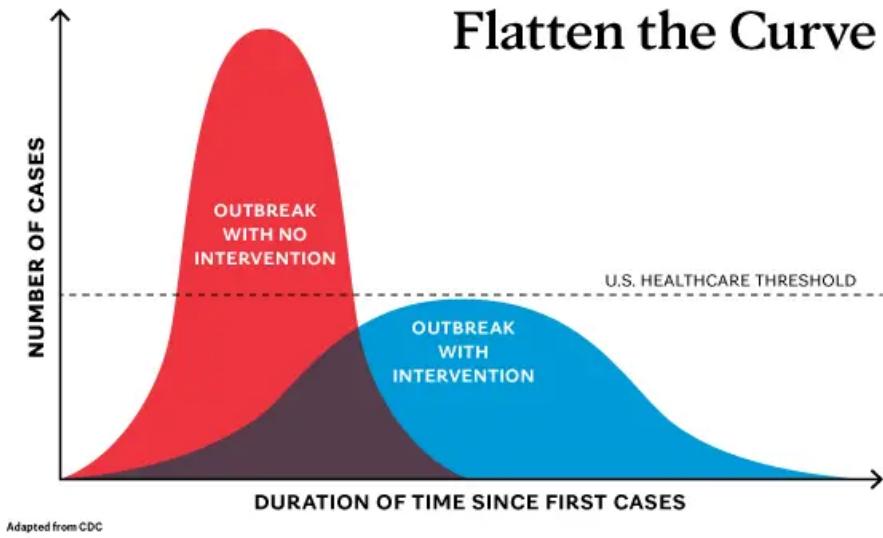


Figure 1.2: Credit to Mayo Clinic News Network

The general problem to minimize the peak infection rate involves L_1 -optimal control, but prior work had been developed in a parameterized or discretized manner [9, 10, 11]. Finding a convex and strongly convergent program to minimize the peak value remained an open problem

CHAPTER 1. INTRODUCTION

during the initial writing of this thesis.

Near the conclusion of my Chateaubriand Fellowship (on July 4, 2022), I attended a **SPOT seminar** (SPOT 75) at ENSEEIHT, Toulouse in which Alain Rapoport discussed his team's work in peak-minimizing Optimal Control Problems (OCPs) [12]. Their formulation involves minimizing a newly added constant-time state variable that always upper-bounds the desired-minimum objective on state trajectories. This concept was the missing link towards providing convex and convergent programs for peak minimizing control. I then used their method to provide further safety quantification techniques (measuring the safety of trajectories by the minimal control effort/data corruption needed to crash).

1.2 Summary

Chapter 2 reviews background information about measure theory, occupation measures, and Linear Matrix Inequality (LMI) approximations to measure LPs.

Past the preliminaries in Chapter 2, the thesis is divided into three parts.

Part 1 extends the peak estimation framework towards safety analysis and systems with uncertainty. Chapter 3 provides an overview of prior work in Ordinary Differential Equation (ODE) peak estimation with infinite-dimensional LPs and introduces a recovery algorithm to attempt extraction of approximately-optimal trajectories. Chapter 4 performs peak estimation on system whose dynamics are affected by bounded-uncertainty processes. Chapter 5 applies peak estimation towards verifying the safety of trajectories with respect to an unsafe set by computing safety margins and distances of closest approach.

Part 2 utilizes robust optimization to solve peak estimation and peak-minimizing-control tasks. Chapter 6 focuses on input-affine dynamics where the input disturbances are restricted to Semidefinite Representable (SDR) sets. Under this structure, the disturbance variables can be eliminated using the theory of robust counterparts, providing tractable SDPs for applications including data-driven peak estimation. Chapter 7 quantifies the safety of trajectories by the required perturbation intensity needed to crash into the unsafe set, and applies this method to analyze data-driven models with respect to their maximal constraint violations. Chapter 8 synthesizes controllers to maximize the distance of closest approach to an unsafe set as trajectories travel from the initial set to the terminal set within a specified time horizon.

Part 3 applies peak estimation methods towards systems with non-ODE dynamical behaviors. Chapter 9 extends the peak estimation framework towards systems with hybrid dynamics.

CHAPTER 1. INTRODUCTION

Chapter 10 introduces probabilistic peak estimation to upper-bound the maximal Value-at-Risk (VaR) of an objective function along Stochastic Differential Equation (SDE) trajectories. Chapter 11 performs peak estimation for Delay Differential Equations (DDEs).

Each chapter between 3-11 begins with an introduction and review of prior art and ends with a conclusion. The content in the chapters may be combined to form a modular framework for peak estimation problems. As an example, Section 5.9.1 (in Chapter 5 about safety) introduces a program to bound the distance of closest approach in which the dynamical system is corrupted by bounded uncertainty (where peak estimation under bounded uncertainty is covered in Chapter 4).

Chapter 12 concludes the thesis and provides an outline for future work.

1.3 Publications

This section lists all disseminated work arising from the thesis. Further detail about the research in this thesis is available at https://jarmill.github.io/projects/peak_project/.

Journal Papers (published)

1. J. Miller, D. Henrion, and M. Sznaier, “Peak Estimation Recovery and Safety Analysis,” *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 1982–1987, 2021 [\[link\]](#)

Journal Papers (conditionally accepted)

1. J. Miller and M. Sznaier, “Bounding the Distance to Unsafe Sets with Convex Optimization,” (Conditionally accepted by IEEE Transactions on Automatic Control in 2022) [\[link\]](#)

Conference Proceedings (published)

1. J. Miller and M. Sznaier, “Bounding the Distance of Closest Approach to Unsafe Sets with Occupation Measures,” in *2022 61st IEEE Conference on Decision and Control (CDC)*, pp. 5008–5013, 2022. [\[link\]](#)
2. J. Miller and M. Sznaier, “Facial Input Decompositions for Robust Peak Estimation under Polyhedral Uncertainty,” *IFAC PapersOnLine*, vol. 55, no. 25, pp. 55–60, 2022. [\[link\]](#). Received IFAC Young Author Award.

CHAPTER 1. INTRODUCTION

3. J. Miller, D. Henrion, M. Sznajer, and M. Korda, “Peak Estimation for Uncertain and Switched Systems,” in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 3222–3228, 2021. [\[link\]](#). Received Outstanding Student Paper Award.

Preprints

1. J. Miller, M. Korda, V. Magron, and M. Sznajer “Peak Estimation of Time Delay Systems using Occupation Measures,” 2023. [\[link\]](#)
2. J. Miller, M. Tacchi, M. Sznajer, and A. Jasour, “Peak Value-at-Risk Estimation for Stochastic Differential Equations using Occupation Measures,” 2023. [\[link\]](#)
3. J. Miller and M. Sznajer, “Peak Estimation of Hybrid Systems with Convex Optimization,” 2023. [\[link\]](#)
4. J. Miller and M. Sznajer “Quantifying the Safety of Trajectories using Peak-Minimizing Control,” 2023. [\[link\]](#)
5. J. Miller and M. Sznajer, “Analysis and Control of Input-Affine Dynamical Systems using Infinite-Dimensional Robust Counterparts,” 2023. [\[link\]](#)

Chapter 2

Preliminaries

This chapter defines notation and reviews background material. The introductory content in this chapter is an extended version of the preliminaries from [13].

2.1 Notation

Let \mathbb{R} be the set of real numbers and \mathbb{R}^n be an n -dimensional real Euclidean space. The set of $m \times n$ matrices with real coefficients is $\mathbb{R}^{m \times n}$, and the set of $n \times n$ symmetric matrices is \mathbb{S}^n . A symmetric matrix $M \in \mathbb{S}^n$ is Positive Semidefinite (PSD) (Positive Definite (PD)) if the quadratic form $x^T M x \geq 0$ ($x^T M x > 0$) for all choices of $x \in \mathbb{R}^n$. The set of n -dimensional symmetric PSD (PD) matrices is denoted by \mathbb{S}_+^n (\mathbb{S}_{++}^n). The notation $M \succeq 0$ will also be used to denote that M is PSD.

Let \mathbb{N} be the set of natural numbers and \mathbb{N}^n be the set of n -dimensional multi-indices. The sequence of natural numbers between a and b (inclusive) is $a..b$. The total degree of a multi-index $\alpha \in \mathbb{N}^n$ is $|\alpha| = \sum_i \alpha_i$. A monomial $\prod_{i=1}^n x_i^{\alpha_i}$ may be expressed in multi-index notation as x^α . The set of polynomials with real coefficients is $\mathbb{R}[x]$, and polynomials $p(x) \in \mathbb{R}[x]$ may be represented as the sum over a finite index set $\mathcal{J} \subset \mathbb{N}^n$ of $p(x) = \sum_{\alpha \in \mathcal{J}} p_\alpha x^\alpha$. The set of polynomials with monomials up to degree $|\alpha| = d$ is $\mathbb{R}[x]_{\leq d}$.

2.2 Measure Theory

This section will introduce concepts in measure theory for use in this thesis. Refer to [14] for a complete reference, and to [15] for a visual introduction to measure theory.

2.2.1 Measures

Let S be a Banach space, and let $\text{Set}(S) = 2^S$ be the set of sets (power set) of a space $S \in \mathbb{R}^n$. A σ -algebra Σ over S is a subset of $\text{Set}(S)$ such that Σ contains S and is closed under countable unions and complements. Examples of σ -algebras over the real line \mathbb{R} include the countable unions of the set of intervals $[a, b]$ with $a, b \in \mathbb{R}$ and when $a, b \in \mathbb{Z}$.

A nonnegative Borel measure $\mu : \Sigma \rightarrow \mathbb{R}_+$ is a function that assigns a size (measure) to each set $A \subset S$ in a σ -algebra Σ under the following rules:

1. $\mu(A) \geq 0 \quad \forall A \in \Sigma$
2. $\mu(\emptyset) = 0$
3. $\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k) \quad \text{sets } A_k \text{ are disjoint.}$

The quantity $\mu(S)$ is known as the mass of μ , and μ is a probability measure if this mass is 1. The Dirac delta $\delta_{s'}$ is a probability measure supported at a single point $s' \in S$. An example of a Dirac Delta with elements (boxes) in $\text{Set}(S)$ are plotted in Figure 2.1. The point s' is the green \bullet -shape in the center of the figure. Boxes which contain s' have a measure of 1 and have solid walls. Boxes that do not contain s' have a measure of 0 and have dotted walls.

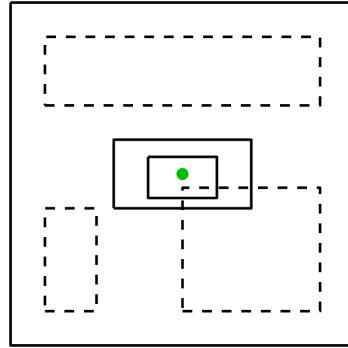


Figure 2.1: Dirac Delta: solid boxes contain x' , dotted boxes miss x'

2.2.2 Pairings and Operators

The set of continuous functions over the Banach space S is denoted as $C(S)$, the set of finite signed Borel measures over S is $\mathcal{M}(S)$, and the set of nonnegative Borel measures over S is $\mathcal{M}_+(S)$. A duality pairing exists between all functions $f \in C(S)$ and measures $\mu \in \mathcal{M}_+(S)$ by Lebesgue

CHAPTER 2. PRELIMINARIES

integration: $\langle f, \mu \rangle = \int_S f(s)d\mu(s)$ when S is compact. The subcone of nonnegative continuous functions over S is $C_+(S) \subset C(S)$, which satisfies $\langle f, \mu \rangle \geq 0 \forall f \in C_+(S), \mu \in \mathcal{M}_+(S)$. The pairing $\langle f, \mu \rangle$ is an inner product between $f \in C_+(S)$ and $\mu \in \mathcal{M}_+(X)$. The sup-norm of a function $f \in C^0(S)$ is $\|f\|_{C^0(S)} = \sup_{s \in S} |f(s)|$. The C^1 norm of a function $f \in C^1(S)$ is $\|f\|_{C^1(S)} = \|f\|_{C^0(S)} + \sum_{i=1}^n \|\partial_{s_i} f\|_{C^0(S)}$. The subcone of continuous functions over S whose first k derivatives are continuous is $C^k(S)$ (with $C(S) = C^0(S)$).

The indicator function of a set $A \subseteq S$ is a function $I_A : S \rightarrow \{0, 1\}$ taking values $I_A(s) = 1$ if $s \in A$ and $I_A(s) = 0$ if $s \notin A$. The measure of a set A with respect to $\mu \in \mathcal{M}_+(S)$ is $\mu(A) = \langle I_A(s), \mu \rangle = \int_A d\mu$ (generalizing the duality pairing to allow for Borel measurable rather than continuous functions). The mass of μ is $\mu(S) = \langle 1, \mu \rangle$, and μ is a probability measure if $\langle 1, \mu \rangle = 1$. The support of μ is the set of all points $s \in S$ such that every open neighborhood N_s of s has $\mu(N_s) > 0$. The Lebesgue measure λ_S over a space S is the volume measure satisfying $\langle f, \lambda_S \rangle = \int_S f(s)ds \forall f \in C(S)$. The Dirac delta $\delta_{s'}$ is a probability measure supported at a single point $s' \in S$, and the duality pairing of any function $f \in C(S)$ with respect to $\delta_{s'}$ is $\langle f(s), \delta_{s'} \rangle = f(s')$. A rank- r atomic measure is a measure μ such there exist scalars $c_i > 0$ and distinct points $s_i \in S$ for $i = 1..r$ such that $\mu = \sum_{i=1}^r c_i \delta_{s_i}$. The atoms of μ are the support points $\{s_i\}_{i=1}^r$.

Let S, Y be spaces and $\mu \in \mathcal{M}_+(S)$, $\nu \in \mathcal{M}_+(Y)$ be measures. The product measure $\mu \otimes \nu$ is the unique measure such that $\forall A \in S, B \in Y : (\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$. The pushforward of a map $Q : S \rightarrow Y$ along a measure $\mu(s)$ is $Q_\# \mu(y)$, which satisfies $\forall f \in C(Y) : \langle f(y), Q_\# \mu(y) \rangle = \langle f(Q(s)), \mu(s) \rangle$. The measure of a set $B \in Y$ is $Q_\# \mu(Y) = \mu(Q^{-1}(Y))$. The projection map $\pi^s : S \times Y \rightarrow S$ preserves only the s -coordinate as $(s, y) \rightarrow s$ and a similar definition holds for π^y . Given a measure $\eta \in \mathcal{M}_+(S \times Y)$, the projection-pushforward $\pi_\#^s \eta$ expresses the s -marginal of η with duality pairing $\forall f \in C(S) : \langle f(s), \pi_\#^s \eta \rangle = \int_{S \times Y} f(s)d\eta(s, y)$. Every linear operator $\mathcal{L} : S \rightarrow Y$ possesses a unique adjoint operator \mathcal{L}^\dagger such that $\langle \mathcal{L}f, \mu \rangle = \langle f, \mathcal{L}^\dagger \mu \rangle, \forall f \in C(S), \mu \in \mathcal{M}_+(S)$.

2.2.3 Absolute Continuity and Domination

Let $\mu, \nu \in \mathcal{M}_+(S)$ be nonnegative Borel measures. The measure ν is absolutely continuous to μ ($\nu \ll \mu$) if, for every $A \in \Sigma$, $\mu(A) = 0$ implies that $\nu(A) = 0$. Equivalently, there exists a unique nonnegative and measurable (density) function $\rho(s)$ such that $\langle f(s), \mu(s) \rangle = \langle f(s)\rho(s), \nu(s) \rangle$ for all $f \in C(S)$. The function ρ is the Radon-Nikodým derivative $\rho = d\nu/d\mu$.

CHAPTER 2. PRELIMINARIES

The measure ν is dominated by μ if $\nu(A) \leq \mu(A)$ for all subsets $A \subset S$ (elements A in a σ -algebra). There exists a unique nonnegative slack measure $\hat{\nu} \in \mathcal{M}_+(S)$ such that $\nu(A) + \hat{\nu}(A) = \mu(A)$, $\forall A \subset S$, which may be equivalently written as $\nu + \hat{\nu} = \mu$. Domination ($\nu \leq \mu$) is a stronger condition than absolute continuity ($\nu \ll \mu$).

A pair of measures $\omega, \nu \in \mathcal{M}_+(S)$ are orthogonal ($\omega \perp \nu$) if $\forall A \in \text{supp}(\omega) : \nu(A) = 0$ and $\forall B \in \text{supp}(\nu) : \omega(B) = 0$. While the sum $\nu + \omega$ dominates ν and ω individually, it does not hold that every μ with $\mu \geq \nu$ and $\mu = \nu + \hat{\nu}$ produces an orthogonal pair $\nu \perp \hat{\nu}$.

If the sets $\{S^i\}_{i=1}^N$ form a partition of S ($\bigcup_{i=1}^N S^i = S$ and $S^i \cap S^{i'} = \emptyset \forall i \neq i'$), then any nonnegative Borel measure $\mu \in \mathcal{M}_+(S)$ may be uniquely split by orthogonality into the sum of measures $\mu_i \in \mathcal{M}_+(S^i) \forall i = 1..N$ with nonoverlapping support such that $\mu = \sum_{i=1}^N \mu_i$.

2.2.4 Signed Measures

A signed measure is a function $\mu : \Sigma \rightarrow \mathbb{R}$ that only satisfies conditions 2 and 3 of a nonnegative measure (sets may have negative measure). The set of signed measures over a space S is $\mathcal{M}(S)$. The Hahn-Jordan decomposition is a unique method to split a signed measure $\mu \in \mathcal{M}$ into the difference of two orthogonal nonnegative measures $\mu = \mu^+ - \mu^-$.

The Total Variation (TV) norm of a signed measure $\mu \in \mathcal{M}(S)$ is

$$\|\mu\|_{TV} = \sup_{v \in C(S)} \langle v, \mu \rangle : -1 \leq v(s) \leq 1 \forall s \in S \quad (2.1a)$$

$$= \inf_{\mu^+, \mu^- \in \mathcal{M}_+(S)} \langle 1, \mu^+ \rangle + \langle 1, \mu^- \rangle : \mu^+ - \mu^- = \mu. \quad (2.1b)$$

2.3 Occupation Measures

Given an interval $[a, b]$ and a continuous curve $s(t)$ where $s : [a, b] \rightarrow S$ and $S \subset \mathbb{R}^n$, the pushforward of the Lebesgue measure on $[a, b]$ through the map $t \rightarrow (t, s(t))$ is called the **occupation measure** associated to $s(t)$ [16].

Assume that there exists a set of states X and a set of initial conditions $X_0 \subseteq X$. Given an initial condition $x_0 \in X_0$, the curve $x(t | x_0)$ corresponds to a trajectory of the following dynamical system starting at time x_0

$$\dot{x}(t) = f(t, x(t)). \quad (2.2)$$

The measure $\mu_{x(\cdot)}$ is the occupation measure associated with $x(t | x_0)$. Given a stopping time $t^* \in [0, T]$, the occupation measure $\mu_{x(\cdot)}$ returns the amount of time the time-indexed trajectory

CHAPTER 2. PRELIMINARIES

$x(t \mid x_0)$ spends in the region $A \times B \subseteq [0, T] \times X$ with

$$\mu_{x(\cdot)}(A \times B \mid x_0) = \int_0^{t^*} I_{A \times B}(t, x(t \mid x_0)) dt. \quad (2.3)$$

The definition in (2.3) induces a pairing rule by integration:

$$\forall \phi \in C([0, T] \times X) : \quad \langle v, \mu \rangle = \int_{t=0}^{t^*} \phi(t, x(t \mid x_0)) dt. \quad (2.4)$$

The occupation measure $\mu_{x(\cdot)}$ may be averaged over a distribution of initial conditions $\mu_0 \in \mathcal{M}_+(X_0)$ to form

$$\mu(A \times B) = \int_{X_0} \mu_{x(\cdot)}(A \times B \mid x_0) d\mu_0(x_0). \quad (2.5)$$

A consequence of (2.5) is that $\langle 1, \mu \rangle \leq T \langle 1, \mu_0 \rangle$.

When $t^* = T$, occupation measures are particular instances of stochastic (Markov) kernels (the conditional distribution $\mu(x \mid t)$ is a probability measure for each fixed t , and the function $\mu(\cdot, B)$ is measurable in t for each $B \in X$).

Figure 2.2 visualizes trajectories in cyan of the dynamical system $x' = -x(x+2)(x-1)$ with one trajectory highlighted in a thick dark blue line. Let $\mu_{x(\cdot)}$ be the occupation measure associated with the thick blue line, and μ be the occupation measure averaged over the uniformly distributed set of initial conditions between $[-4, 4]$. No trajectory passes through the green box the top-right, so the measures of $\mu(\text{green})$ and $\mu_{x(\cdot)}(\text{green})$ are both zero. The thick blue trajectory passes through the black box on the bottom, so $\mu(\text{black})$ and $\mu_{x(\cdot)}(\text{black})$ are both nonzero. The red box on the left does intersect trajectories starting between $[-4, 4]$ but does not contain the blue trajectory, so $\mu(\text{red})$ is nonzero and $\mu_{x(\cdot)}(\text{red})$ is zero.

For a test function $v(t, x) \in C^1([0, T] \times X)$, the Lie derivative operator \mathcal{L}_f is defined

$$\mathcal{L}_f v(t, x) = \partial_t v(t, x) + \nabla_x v(t, x)^T f(t, x). \quad (2.6)$$

The measure $\mu_p \in \mathcal{M}_+([0, T] \times X)$ is a free-terminal-time distribution of times and states. This three measures μ_0, μ_p, μ are linked together by the linear Liouville Equation

$$\langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle \mathcal{L}_f v(t, x), \mu \rangle \quad \forall v \in C^1([0, T] \times X). \quad (2.7)$$

Equation (2.7) may be equivalently expressed in shorthand notation (abstracting out the $\forall v$ imposition) as

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu. \quad (2.8)$$

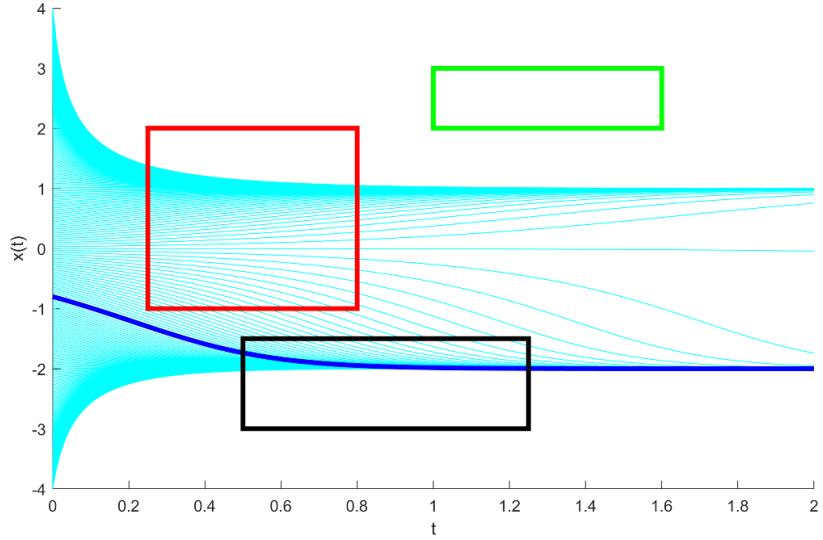


Figure 2.2: Trajectories of $x' = -x(x + 2)(x - 1)$ and boxes

The operator \mathcal{L}_f^\dagger is the adjoint of \mathcal{L}_f such that $\langle \mathcal{L}_f v, \mu \rangle = \langle v, \mathcal{L}_f^\dagger \mu \rangle$ for any $v(t, x) \in C^1([0, T] \times X)$. Every trajectory $x(t \mid x_0)$ with $x_0 \in X_0$ and $t^* \in [0, T]$ induces a measure representation following (2.7) and (2.8) with $\mu_0 = \delta_{x=x_0}$, $\mu = \mu_{x(\cdot)}$, and $\mu_p = \delta_{t=t^*} \otimes \delta_{x=x(t^*|x_0)}$.

2.4 Moment-SOS Hierarchy

The standard form for a measure Linear Program (LP) with variable $\mu \in \mathcal{M}_+(X)$ involving a cost function $p \in C(X)$ and a (possibly infinite) set of affine constraints $\langle a_j, \mu \rangle = b_j$ with $b_j \in \mathbb{R}$ and $a_j \in C(X)$ for $j = 1..J_{max}$ is

$$p^* = \sup_{\mu \in \mathcal{M}_+(X)} \langle p, \mu \rangle \quad (2.9a)$$

$$\langle a_j(x), \mu \rangle = b_j \quad \forall j = 1..J_{max}. \quad (2.9b)$$

The dual problem to Program (2.9) with dual variables $v_j \in \mathbb{R} : \forall j = 1..m$ is,

$$d^* = \inf_{v \in \mathbb{R}^m} \sum_j b_j v_j \quad (2.10a)$$

$$p(x) - \sum_j a_j(x)v_j \geq 0 \quad \forall x \in X. \quad (2.10b)$$

The objectives in (2.9) and (2.10) will match ($p^* = d^*$ strong duality) if p^* is finite and if the mapping $\mu \rightarrow \{\langle a_j(x), \mu \rangle\}_{j=1}^m$ is closed in the weak-* topology (Theorem 3.10 in [17]).

CHAPTER 2. PRELIMINARIES

When $p(x)$ and all $a_j(x)$ are polynomial, Equation (2.10b) is a polynomial nonnegativity constraint.

The restriction that a polynomial $q(x) \in \mathbb{R}[x]$ is nonnegative over \mathbb{R}^n may be strengthened to finding a set of polynomials $\{q_i(x)\}$ such that $q(x) = \sum_i q_i(x)^2$. The polynomials $\{q_i(x)\}$ are a Sum of Squares (SOS) certificate of nonnegativity of $q(x)$, given that the square of a real quantity $q_i(x)$ at each i and x is nonnegative. The set of SOS polynomials in indeterminate quantities x is expressed as $\Sigma[x]$, with a maximal-degree- d subset of $\Sigma[x]_{\leq d}$. A polynomial $p(x)$ is SOS ($p(x) \in \Sigma[x]$) iff there exists a finite integer s , a polynomial vector $v(x) \in \mathbb{R}[x]^s$, and a PSD matrix $Q \in \mathbb{S}_+^s$, such that $p(x) = v(x)^T Q v(x)$. SOS polynomials are nonnegative over \mathbb{R}^n .

A Basic Semialgebraic (BSA) set $\mathbb{K} = \{x \mid g_i(x) \geq 0, i = 1..N_c\}$ is a set formed by a finite set of bounded-degree polynomial constraints.

The quadratic module $Q[g]$ formed by the constraints describing the BSA set $\mathbb{K} = \{x \mid g_i(x) \geq 0, i = 1..N_c\}$ is the set of polynomials:

$$Q[g] = \left\{ \sigma_0(x) + \sum_{i=1}^{N_c} \sigma_i(x) g_i(x) \right\}, \quad (2.11)$$

such that the multipliers σ are SOS:

$$\sigma_i(x) \in \Sigma[x] \quad \forall i = 0..N_c. \quad (2.12)$$

The BSA set \mathbb{K} is compact if there exists a constant $0 \leq R < \infty$ such that \mathbb{K} is contained in the ball $R \leq \|x\|_2^2$. \mathbb{K} satisfies the Archimedean property if the polynomial $R - \|x\|_2^2$ is a member of $Q[g]$. The Archimedean property is stronger than compactness [18], and compact sets may be rendered Archimedean by adding a redundant ball constraint $R - \|x\|_2^2 \geq 0$ to the list of constraints describing in \mathbb{K} (though finding such an R may be difficult). When \mathbb{K} is Archimedean, every polynomial satisfying $p(x) > 0, \forall x \in \mathbb{K}$ has a representation (Putinar's Positivstellensatz [19]):

$$\begin{aligned} p(x) &= \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) \\ \sigma_0(x) &\in \Sigma[x] \quad \sigma_i(x) \in \Sigma[x]. \end{aligned} \quad (2.13)$$

The Weighted Sum of Squares (WSOS) set $\Sigma[\mathbb{K}]$ is the set of polynomials that admit a positivity certificate over \mathbb{K} from (2.13). Its maximal degree- d subset is $\Sigma[\mathbb{K}]_{\leq d}$. Given a multi-index $\alpha \in \mathbb{N}^n$, the α -moment of a measure $\mu \in \mathcal{M}_+(X)$ is $\mathbf{m}_\alpha = \langle x^\alpha, \mu \rangle$.

A measure μ is a representing measure for a moment sequence $\tilde{\mathbf{m}}$ if $\tilde{\mathbf{m}}_\alpha = \langle x^\alpha, \mu \rangle \forall \alpha \in \mathbb{N}^n$. The measure μ is additionally moment-determinate if μ is the unique representing measure associated with $\tilde{\mathbf{m}}$ (Def. 1.4 of [20]). An infinite moment matrix $\mathbb{M}[\mathbf{m}]_{\alpha,\beta} = \mathbf{m}_{\alpha+\beta}$ indexed by

CHAPTER 2. PRELIMINARIES

monomials $\alpha, \beta \in \mathbb{N}^n$ may be constructed from the moment sequence \mathbf{m} . One sufficient condition for a sequence \mathbf{m} to be moment-determinate on the compact set $[-a, a]^n$ for some finite $R, a > 0$ is that $\mathbb{M}[\mathbf{m}] \succeq 0$ and $|\mathbf{m}_\alpha| \leq Ra^{|\alpha|} \forall \alpha \in \mathbb{N}^n$ (Lemma 1.4 of [20]).

The degree- d moment matrix $\mathbb{M}_d[\mathbf{m}]$ of size $\binom{n+d}{d}$ is the submatrix of $\mathbb{M}[\mathbf{m}]$ where the indices $\mathbb{M}_d[\mathbf{m}]_{\alpha,\beta}$ have total degree bounded by $0 \leq |\alpha|, |\beta| \leq d$. Given a polynomial $g(x) \in \mathbb{R}[x]$, the localizing matrix associated with g is a square infinite-dimensional symmetric matrix with entries $\mathbb{M}[g\mathbf{m}]_{\alpha,\beta} = \sum_{\gamma \in \mathbb{N}^n} g_\gamma \mathbf{m}_{\alpha+\beta+\gamma}$. A moment sequence \mathbf{m} has a representing measure $\mu \in \mathcal{M}_+(\mathbb{K})$ if there exists μ supported in \mathbb{K} such that $\mathbf{m}_\alpha = \langle x^\alpha, \mu \rangle \forall \alpha \in \mathbb{N}^n$. The Linear Matrix Inequality (LMI) conditions that $\mathbb{M}[\mathbf{m}] \succeq 0$ and $\mathbb{M}[g_i \mathbf{m}] \succeq 0 \forall i = 1..N_c$ are necessary to guarantee the existence of a representing measure associated with \mathbf{m} . The moment matrix $\mathbb{M}[\mathbf{m}]$ is a localizing matrix with the function $g = 1$. These LMI conditions are sufficient if the set \mathbb{K} is Archimedean, and all compact sets may be rendered Archimedean through the application of a redundant ball constraint [19].

Assume that each polynomial $g_i(x)$ in the constraints of \mathbb{K} induces a degree $d_i = \lceil \deg g_i / 2 \rceil$. We define a block-diagonal matrix $\mathbb{M}_d[\mathbb{K}\mathbf{m}]$ containing the moment and all localizing matrices as,

$$\text{diag}(\mathbb{M}_d[\mathbf{m}], \{\mathbb{M}_{d-d_i}(g_i \mathbf{m}) \forall i = 1..N_c\}). \quad (2.14)$$

The degree- d moment relaxation of Problem (2.9) with variables $y \in \mathbb{R}^{\binom{n+2d}{2d}}$ is

$$p_d^* = \max_{\mathbf{m}} \sum_{\alpha} p_{\alpha} \mathbf{m}_{\alpha}, \quad \mathbb{M}_d[\mathbb{K}\mathbf{m}] \succeq 0 \quad (2.15a)$$

$$\sum_{\alpha} a_{j\alpha} \mathbf{m}_{\alpha} = b_j \quad \forall j = 1..m. \quad (2.15b)$$

The bound $p_d^* \geq p^*$ is an upper bound for the infinite-dimensional measure LP. The decreasing sequence of upper bounds $p_d^* \geq p_{d+1}^* \geq \dots \geq p^*$ is convergent to p^* as $d \rightarrow \infty$ if \mathbb{K} is Archimedean. The dual semidefinite program to (2.15a) is the degree- d SOS relaxation of (2.10):

$$d_d^* = \min_{v \in \mathbb{R}^m} \sum_j b_j v_j \quad (2.16a)$$

$$p(x) - \sum_j a_j(x) v_j = \sigma_0(x) + \sum_k \sigma_i(x) g_i(x) \quad (2.16b)$$

$$\sigma(x) \in \Sigma[x]_{\leq 2d} \quad (2.16c)$$

$$\sigma_i(x) \in \Sigma[x]_{\leq 2d - \lceil \deg g_i / 2 \rceil} \quad \forall i \in 1..N_c. \quad (2.16d)$$

We use the convention that the degree- d SOS tightening of (2.16) involves polynomials of maximal degree $2d$. When the moment sequence \mathbf{m}_α is bounded ($|\mathbf{m}_\alpha| < \infty \forall |\alpha| \leq 2d$) and there exists an interior point of the affine measure constraints in (2.9b), then the finite-dimensional truncations

CHAPTER 2. PRELIMINARIES

(2.15a) and (2.16) will also satisfy strong duality $p_k^* = d_k^*$ at each degree k (by arguments from Appendix D/Theorem 4 of [21] using Theorem 5 of [22], also applied in Corollary 8 of [23]). The sequence of upper bounds (outer approximations) $p_d^* \geq p_{d+1}^* \geq \dots$ computed from Semidefinite Programs (SDPs) is called the Moment-SOS hierarchy.

Let \mathbf{m} be a moment sequence with $\mathbb{M}_d[\mathbf{m}] \succeq 0$. The sequence \mathbf{m} has a *flat extension* if $\text{rank}(\mathbb{M})_d[\mathbf{m}] = \text{rank}(\mathbb{M})_{d-1}[\mathbf{m}]$, which implies that there exists an atomic representing measure for \mathbf{m} with $\text{rank}(\mathbb{M})_d[\mathbf{m}]$ atoms [24, 25]. The sequence \mathbf{m} has a *flat extension* for the \mathbb{K} -constrained moment problem if $\mathbf{m}_d[\mathbb{K}\mathbf{m}] \succeq 0$ and $\text{rank}(\mathbb{M})_d[\mathbf{m}] = \text{rank}(\mathbb{M})_{d-\max_i d_i}[\mathbf{m}]$, and has an atomic representing measure with $\text{rank}(\mathbb{M})_d[\mathbf{m}]$ [26] [27, Theorem 3.11]. Flat extensions have applications in certification of global optimality for polynomial optimization problems.

Part 1: Peak Estimation and Safety Analysis

Chapter 3

Peak Estimation and Recovery

3.1 Introduction

The behavior of dynamical systems may be analyzed by bounding extreme values of state functions along trajectories. For a system with dynamics governed by an Ordinary Differential Equation (ODE) $\dot{x} = f(t, x)$ with continuous f , let $x(t \mid x_0)$ denote a trajectory starting from an initial point x_0 . The problem of finding the maximum value of a function $p(x)$ for trajectories starting from a set X_0 evolving over the time interval $[0, T]$ is

$$\begin{aligned} P^* &= \sup_{t, x_0 \in X_0} p(x(t)) \\ \dot{x}(t) &= f(t, x), \quad t \in [0, T]. \end{aligned} \tag{3.1}$$

The goal of peak estimation is to approximate sharp upper bounds to P^* . It is also desired to recover the near-optimal trajectories that achieve $p(x(t \mid x_0)) \approx P^*$ for some time $t \in [0, T]$. Lower bounds to P^* can be found by sampling an initial point $x_0 \in X_0$ and finding the maximum value of $p(x)$ along $x(t \mid x_0)$, but generating a sampled lower bound that is close to P^* is difficult. Upper bounds of P^* are universal properties of all trajectories, and P^* may be sandwiched between discovered lower and upper bounds. Peak estimation may be infinite-time if $T = \infty$.

Problem (3.1) was cast into an infinite-dimensional LP of occupation measures in the context of optimal stopping problems of a martingale in [5], and the bound P^* was approximated by discretization with finite-dimensional LPs. The infinite-dimensional LP in [5] is an extension to the stochastic setting of the deterministic optimal control formulation in [7] with a state cost instead of a running cost. A survey of infinite-dimensional LP methods is available at [28], and LPs in occupation measures may also be solved through the moment-SOS hierarchy of SDP [16]. More recently, an

CHAPTER 3. PEAK ESTIMATION AND RECOVERY

auxiliary function approach was developed to find a convergent sequence of upper bounds to P^* by SOS methods [6]. The infinite-dimensional LP in [6] that is truncated into an SOS program is dual to LMI relaxations of the infinite-dimensional LP in [5]. The optimal trajectories that achieve P^* are localized into a sublevel set of the solved auxiliary function, and may be approximated through adjoint optimization [29, 6]. The constraints of the SOS programs in [6] are dual to LMIs in moments of occupation measures [16, 21].

This chapter reviews measure-LP formulations for peak estimation problems, and introduces a recovery algorithm to approximate optimizing trajectories. If moment matrices of the LMI solution satisfy an approximate rank constraint, an attempt may be made to extract near-optimal trajectories through an atom extraction procedure (Cholesky decomposition) [30]. Other methods for trajectory extraction require additional postprocessing (by solving optimization problems) after the LMI solution is computed. One of these other recovery techniques include adjoint optimization within an intersection of sublevel sets [6]. Another recovery technique is the application of Christoffel-Darboux kernels to isolate the support of the occupation measure [31] from its approximate moments (appropriate moments of $\mathbb{M}_{d'_k}(y^k)$).

This chapter is organized as follows: Section 3.2 reviews LP formulations for peak estimation problems. Section 3.3 posits a recovery algorithm to extract near-optimal trajectories. Section 3.4 presents a set of numerical examples with successful extraction by the recovery algorithm. Section 3.5 extends the rank-based recovery algorithm to peak estimation problems over global attractors. The recovery section of this chapter is from [32], and was coauthored by Didier Henrion and Mario Sznaier.

3.2 Peak Estimation Programs

Peak estimation problems can be bounded by an infinite-dimensional LP in measures by defining a peak measure $\mu_p \in \mathcal{M}_+([0, T] \times X)$, which generalizes $\delta_T \otimes \mu_T$ with free terminal time.

CHAPTER 3. PEAK ESTIMATION AND RECOVERY

Eq. (9) from [5] with variables (μ_0, μ, μ_p) can be restated as

$$p^* = \sup \langle p(x), \mu_p \rangle \quad (3.2a)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \quad (3.2b)$$

$$\mu_0(X_0) = 1 \quad (3.2c)$$

$$\mu, \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (3.2d)$$

$$\mu_0 \in \mathcal{M}_+(X_0). \quad (3.2e)$$

The probability measure μ_0 in (3.2c) is distributed over initial conditions. By Liouville's Equation (3.2b), μ_p is a probability measure over points in time and space:

$$\langle 1, \mu_p \rangle = \langle 1, \delta_0 \otimes \mu_0 \rangle + \langle \mathcal{L}_f(1), \mu \rangle = 1 + 0 = 1. \quad (3.3)$$

Another consequence of the Liouville equation is

$$\langle t, \mu_p \rangle = \langle t, \delta_0 \otimes \mu_0 \rangle + \langle \mathcal{L}_f(t), \mu \rangle = 0 + \langle 1, \mu \rangle = \langle 1, \mu \rangle. \quad (3.4)$$

The dual problem to (3.2) with variables $(v(t, x), \gamma)$ is

$$d^* = \inf_{\gamma \in \mathbb{R}} \gamma \quad (3.5a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (3.5b)$$

$$\mathcal{L}_f v(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (3.5c)$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (3.5d)$$

$$v \in C^1([0, T] \times X), \quad (3.5e)$$

and is formulated in Eq. 2.5 and 2.6 of [6]. The auxiliary function $v(t, x)$ and scalar γ are dual variables for constraints (3.2b) and (3.2c) respectively [6]. If (v, γ) solves to (3.5), then the sublevel set $\{(t, x) \mid v(t, x) \leq \gamma\}$ contains all trajectories starting from X_0 .

The solution $p^* = d^* \geq P^*$ is an upper bound for the true peak in (3.1). Strong duality holds with $p^* = d^*$ when $\{[0, T], X, X_0\}$ are all compact [7]. The solution p^* is approximately equal to P^* for compact $[0, T] \times X$ and locally Lipschitz dynamics (Theorem 2.1 of [7], 2.5 of [6]), and often $p^* = P^*$. The objective values $d^* = P^*$ are tight if the function $v(t, x)$ in (3.5) is allowed to be discontinuous [6].

The work in [5] estimates (3.2) by discretizing the infinite-dimensional LP (sec. 4.1) or forming a Markov chain (sec. 4.2). The work in [6] finds a convergent sequence of upper bounds through an SOS relaxation (Eq. 4.4-4.7).

3.3 Recovery

This section presents an algorithm to attempt extraction of optimal trajectories if p^* is reached at a finite number of R points. This recovery algorithm was first presented in [32].

3.3.1 Optimal Trajectories and Measures

Each of the R solution trajectories to Problem (3.1) that achieves P^* may be encoded by a triple (x_0^r, t_p^r, x_p^r) satisfying $P^* = p(x_p^r) = p(x(t_p^r | x_0^r))$ for $r = 1..R$. A trajectory $x(t | x_0)$ in which P^* is reached multiple times is separated into triples for each attainment.

Let the triple (x_0, t_p, x_p) be a solution to Problem (3.1). The probability measures $\mu_0 = \delta_{x_0}$, $\mu_p = \delta_{t_p} \otimes \delta_{x_p}$, and μ defined by Eq. (2.3) with an endpoint t_p instead of T , satisfy constraints (3.2b)-(3.2e) with an objective value of $\langle p, \mu_p \rangle = P^*$ (where μ is supported between $(0, x_0)$ and (t_p, x_p)). For the general case where P^* is reached at multiple triples (x_0^r, t_p^r, x_p^r) , the measures $\mu_0 = \sum_{r=1}^R w_r \delta_{x_0^r}$, $\mu_p = \sum_{r=1}^R w_r (\delta_{t_p^r} \otimes \delta_{x_p^r})$, and $\mu = \sum_{r=1}^R w_r \mu^r$ are feasible solutions to (3.2b)-(3.2e) for all weights $w \in \mathbb{R}_+^R$ with $\mathbf{1}^T w = 1$ (convex combinations). If $p^* = P^*$, optimal trajectories may be recovered from the support of μ_0 and μ_p solving (3.2).

3.3.2 LMI Formulation

Assume that the measures μ_0, μ, μ_p from (3.2) have moment sequences of $\mathbf{m}^0, \mathbf{m}, \mathbf{m}^p$ up to degree $2d$. Liouville's equation in (3.2b) implies that the following linear relation holds for each test function $v(t, x) = x^\alpha t^\beta$:

$$\langle x^\alpha, \mu_0 \rangle \delta_{\beta 0} + \langle \mathcal{L}_f(x^\alpha t^\beta), \mu \rangle - \langle x^\alpha t^\beta, \mu_p \rangle = 0. \quad (3.6)$$

The expression $\delta_{\beta 0}$ is the Kronecker Delta taking a value $\delta_{\beta 0} = 1$ when $\beta = 0$ and $\delta_{\beta 0} = 0$ when $\beta \neq 0$.

Define $\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^p)$ as the relation in moment sequences from (3.6) for each test function $x^\alpha t^\beta$, and the dynamics degree \tilde{d} as $d + \lfloor \deg f / 2 \rfloor$ for a given degree d . Assuming that p and the entries of f are given polynomials and

$$X = \{x \mid g_k(x) \geq 0, \forall k = 1..N_c\} \quad (3.7a)$$

$$X_0 = \{x \mid g_{0k}(x) \geq 0, \forall k = 1..N_c^0\} \quad (3.7b)$$

CHAPTER 3. PEAK ESTIMATION AND RECOVERY

are compact basic semialgebraic sets, the degree- d LMI relaxation of (3.2) with variables $(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^p)$ is

$$p_d^* = \max \quad \sum_{\alpha} p_{\alpha} y_{\alpha}^p. \quad (3.8a)$$

$$\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^p) = 0 \quad \forall (\alpha, \beta) \in \mathbb{N}_{\leq 2d}^{m+1} \quad (3.8b)$$

$$\mathbf{m}_0^0 = 1 \quad (3.8c)$$

$$\mathbb{M}_d(X_0 \mathbf{m}^0) \succeq 0 \quad (3.8d)$$

$$\mathbb{M}_d(([0, T] \times X) \mathbf{m}^p) \succeq 0 \quad (3.8e)$$

$$\mathbb{M}_{\tilde{d}}(([0, T] \times X) \mathbf{m}) \succeq 0. \quad (3.8f)$$

Program (3.8) is dual to the degree- d SOS program in [6]. Localizing matrix constraints (3.8d)-(3.8f) enforce the measure support constraints in (3.2d)-(3.2e).

3.3.3 Recovery Algorithm

A solution to (3.8) at degree d will yield an upper bound $p_d^* \geq p^*$. There exist atomic representing measures μ_0, μ_p whose measures agree with the moment sequences up to order $2d$ if the moment matrices $\mathbb{M}_d(\mathbf{m}^0)$ and $\mathbb{M}_d(\mathbf{m}^p)$ are rank-deficient and have a flat extension. These representing measures may not necessarily solve (3.2), as there may not exist a μ supported on the graph of optimal trajectories with moments in $\mathbb{M}_d(\mathbf{m})$.

The atoms of $\mathbb{M}_d(\mathbf{m}^0)$ and $\mathbb{M}_d(\mathbf{m}^p)$ with extraction by [33] (or reading $\mathbf{m}^0, \mathbf{m}^p$ if rank-1, which automatically implies existence of a flat extension) are candidates for optimal triples (x_0^r, t_p^r, x_p^r) . Evaluating $p(x)$ along a sampled trajectory starting at a feasible atom $x_0^r \in X_0$ from $\mathbb{M}_d(\mathbf{m}^0)$ will yield a lower bound p_d^r such that $p_d^r \leq p^* \leq p_d^*$. If the lower and upper bound are sufficiently close together, then the trajectory starting at x_0^r is approximately optimal. Algorithm 1 describes the forward trajectory recovery algorithm. An alternative approach could take atoms from μ_p with $p_d^* - p(x_p^r) \leq \epsilon$, running f backwards from x_p^r for time t_p^r , and observing if the destination point is a member of X_0 .

This process assumes that the peak estimation problem takes an optimal value at a finite set of points, which in practice is not very restrictive. The rank-recovery process requires low rank moment matrices, is sensitive to numerical conditioning in the monomial basis as d increases, and may not always succeed (e.g., $p^* - P^* > \epsilon$). A dynamical system that possesses symmetry under action by a continuous group may have a set of optimal solutions with dimension greater than zero.

The super-resolution procedure in [34] may be used to approximate the starting points x_0^r and peak points x_p^r in case their moment matrices are not sufficiently low rank.

After an appropriate symmetry reduction [35], the moment matrices $\mathbb{M}_d(\mathbf{m}^0)$ and $\mathbb{M}_d(\mathbf{m}^p)$ may contain a discrete set of atoms, each corresponding to orbit representatives of the symmetry group.

Algorithm 1: Trajectory recovery

Input : Sets X_0 , X , dynamics f , cost p , max. time T , initial degree d_0 , tolerance ϵ

Output : Near-Optimal Trajectories OPT

degree $d = d_0$, optimal triples $OPT = \emptyset$

Loop

Solve (3.8) at degree d for $(p_d^*, \mathbb{M}_d(\mathbf{m}^0))$

if $\mathbb{M}_d(\mathbf{m}^0)$ has a flat extension **then**

for atoms x_0^r in $\mathbb{M}_d(\mathbf{m}^0)$ by [33] **do**

 Simulate $x(t | x_0^r)$

 Find $p_d^r = \max_{t \in [0, T]} p(x(t | x_0^r))$

 Find t_p^r, x_p^r on traj. with $p_d^r = p(x_p^r)$

if $p_d^* - p_d^r < \epsilon$ **then**

 Append (x_0^r, t_p^r, x_p^r) to OPT

end

return OPT **if** $OPT \neq \emptyset$

$d \leftarrow d + 1$

EndLoop

3.4 Recovery Examples

3.4.1 Flow System

A persistent example throughout this thesis will be the Flow system of [36]:

$$\dot{x} = \begin{bmatrix} x_2 \\ -x_1 - x_2 + \frac{1}{3}x_1^3 \end{bmatrix}. \quad (3.9)$$

Figure 3.1 plots trajectories of the flow system in cyan for times $t \in [0, 5]$, starting from the initial set $X_0 = \{x \mid (x_1 - 1.5)^2 + x_2 \leq 0.4^2\}$ in the black circle. The minimum value of x_2 along

these trajectories is $\min x_2 \approx -0.5734$. The optimizing trajectory is shown in dark blue, starting at the blue circle $x_0^* = (1.4889, -0.3998)$, and reaching optimality at $x_p^* = (0.6767, -0.5734)$ in time $t_p^* = .6627$.

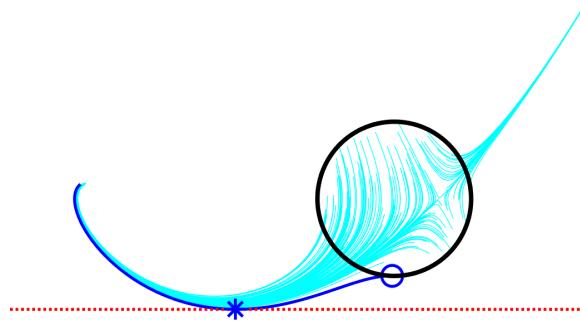


Figure 3.1: Minimizing x_2 along Flow system (3.9)

3.4.2 Symmetric Attractor

Example 4.1 from [6] is the following system with a central symmetry and two stable attractors:

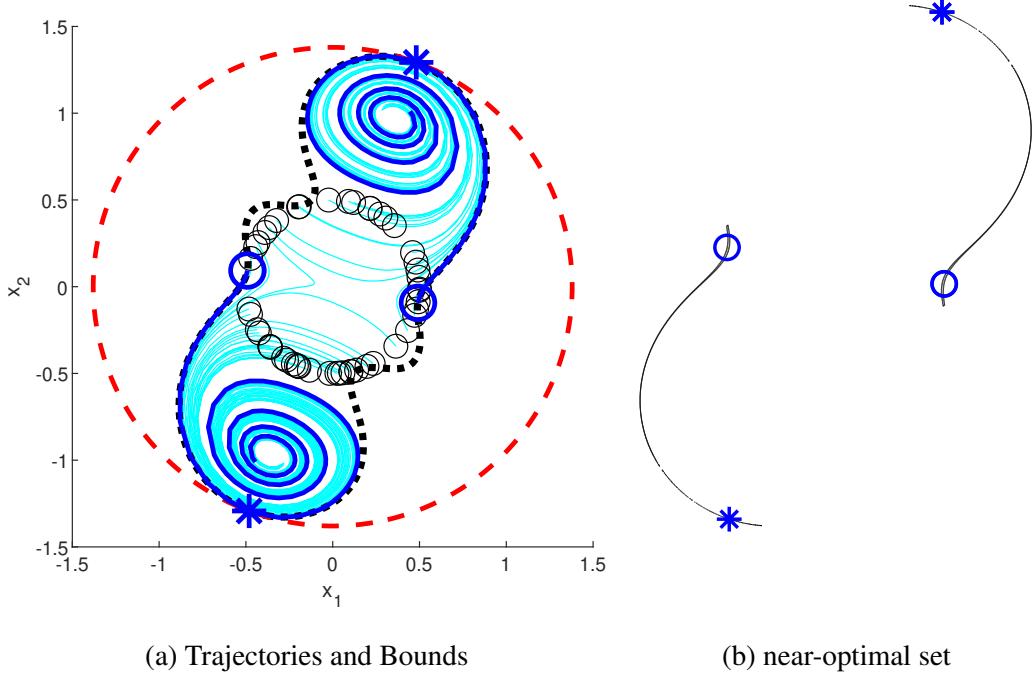
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0.2x_1 + x_2 - x_2(x_1^2 + x_2^2) \\ -0.4x_1 + x_1(x_1^2 + x_2^2) \end{bmatrix}. \quad (3.10)$$

For the infinite-horizon problem (without variable t) of maximizing $\|x\|_2^2$ starting at $X_0 = \{x \mid \|x\|_2^2 = 0.5\}$, $X = [-2, 2]^2$, (3.8) finds a bound $p_7^* = 1.90318$. The solved $\mathbb{M}_7(y_0), \mathbb{M}_7(y_p)$ are rank-2 up to a tolerance of 3×10^{-4} . When using Alg. 1, p_7^* is within 0.005 of the sampled result p_7^r of each atom.

Fig. 3.2a plots the optimal trajectory in dark blue and randomly sampled trajectories in cyan along with the level set $p(x) = p_7^*$ in the red dashed line. The black dashed curve is the level set $\{x \mid v(x) = 0\}$. Fig. 3.2b compares the extracted $x_0^* \approx \pm(0.491, -0.093)$ (blue circles) and $x_p^* \approx \pm(0.481, 1.293)$ (blue stars) against a sublevel-set approximation to locations of optimal trajectories and their initial conditions ([6] Sec. 3: $\{x \mid 0 \leq -v(x) + p_7^* \leq 0.002, 0 \leq -\mathcal{L}_f v(x) \leq 0.004\}$).

3.4.3 Frictioned Pendulum

Another example of recovery is in finding the maximum height of a unit pendulum with friction. Pendulum dynamics with angle θ and angular velocity $\dot{\theta} = \omega$ are $\dot{\omega} = -\sin \theta - 0.1\omega$. The


 Figure 3.2: Maximize $\|x\|_2^2$ along (3.10)

initial set is $X_0 = \{\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], \omega = [-1, 1]\}$. The trigonometric expression is reformulated into polynomial dynamics in terms of $c = \cos \theta, s = \sin \theta$ satisfying $c^2 + s^2 = 1$:

$$\begin{bmatrix} \dot{c} \\ \dot{s} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} -s\omega \\ c\omega \\ -s - 0.1\omega \end{bmatrix}. \quad (3.11)$$

The $d = 4$ LMI relaxation in Section 3.2 produces an upper bound on pendulum height of $h_4^* = 1 - \cos \theta^* \approx 1.4682$ over $t \in [0, \infty)$, with results provided in Figure 3.3. The initial points generating the optimal trajectory are $x_0 = [\theta_0, \omega_0] = \pm[\frac{\pi}{2}, 1]$, and the peak is achieved at $x_p = \pm[2.058, 0]$ (swing angle of $\theta^* \approx 117.92^\circ$). Figure 3.3 displays the upper bound as a red plane. The two optimal trajectories generated by Algorithm (1) are marked by thick blue curves, originating from x_0^* (circles) and reaching the maximum height x_p^* (stars). The black contour is the invariant set $\{x \mid v(x) = h_4^*\}$.

3.5 Global Attractor Peak Recovery

With minor modifications, Algorithm (1) may be used to recover points on a global attractor that maximize a function $p(x)$.

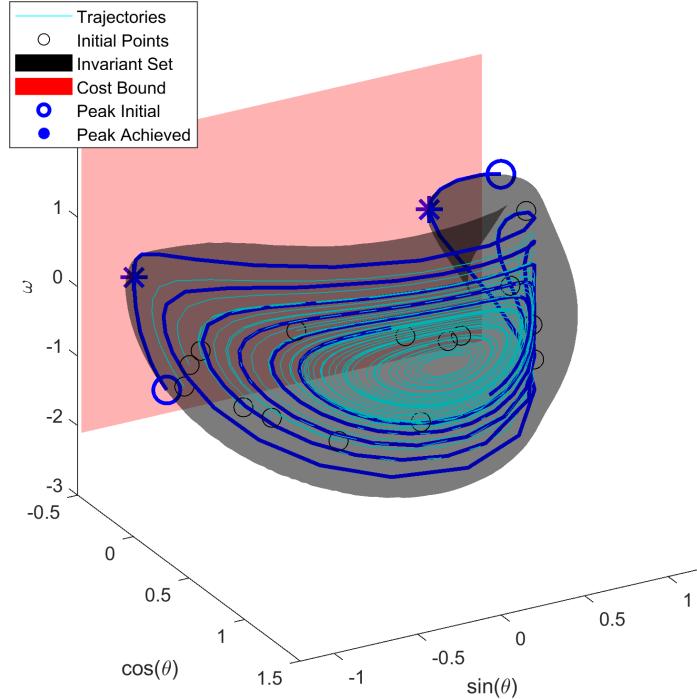


Figure 3.3: Maximize $h = 1 - \cos \theta$ along pendulum with friction

3.5.1 Global Attractor Background

This subsection follows the background of [37]. A set X is forward invariant with respect to dynamics $\dot{x} = f(t, x)$ if the initial condition $x_0 \in X$ implies that $x(t | x_0) \in \mathcal{S}$, $\forall t \in [0, \infty)$. The maximum positively invariant (MPI) set M_+ is the largest forward-invariant subset of the set X .

A global attractor $\mathcal{A} \subset X$ is the smallest compact set satisfying the property

$$\lim_{t \rightarrow \infty} \text{dist}(x(t, x_0), \mathcal{A}) = 0, \quad \forall x_0 \in M_+, \quad (3.12)$$

where the dist is the minimal distance (for some metric) from a point in X to the set \mathcal{A} . The global attractor set is forward and backward invariant in time [38].

3.5.2 Global Attractor Program

The problem of peak estimation over global attractors may be posed as

$$P^* = \max_{x \in \mathcal{A}} p(x) \quad \dot{x} = f(x). \quad (3.13)$$

Problem (3.13) has been upper-bounded by a sum-of-squares program enforcing that \mathcal{A} is an attractive set in [39]. [37] proposes an infinite-dimensional linear program in discounted

CHAPTER 3. PEAK ESTIMATION AND RECOVERY

occupation measures to provide an outer approximation for \mathcal{A} by the mechanisms of reachability set estimation. For a discount factor $\alpha > 0$, a subset $C \subset X$, and a point $x_0 \in X$, define the α -discounted occupation measure μ as

$$\mu(C) = \int_0^\infty e^{-\alpha t} I_C(x(t | x_0)) dt. \quad (3.14)$$

The α -discounting is used to enforce that μ has bounded mass as $t \rightarrow \infty$. The average α -discounted occupation measure may be formed with respect to an initial measure $\mu_0 \in \mathcal{M}_+(X)$

$$\mu(C) = \int_X \int_0^\infty e^{-\alpha t} I_C(x(t | x_0)) dt d\mu_0(x_0). \quad (3.15)$$

Liouville's equation may be formed for trajectories on a global attractor. For a measure $\mu_0 \in \mathcal{M}_+(X)$ distributed over points and an α -discounted occupation measure $\mu \in \mathcal{M}_+(X)$, a Liouville equation holds for all test functions $v(x) \in C^1(X)$:

$$\beta \langle v, \mu \rangle - \langle \mathcal{L}_f v, \mu \rangle = \langle v, \mu_0 \rangle. \quad (3.16)$$

3.5.2.1 Measure Program

Equation 6 of [37] may be modified for use in peak estimation (μ_0 has unit mass, no Lebesgue measure, objective p), forming the measure program

$$p^* = \sup \langle p(x), \mu_0 \rangle \quad (3.17a)$$

$$\alpha \mu_+ - \mathcal{L}_f^\dagger \mu_+ = \mu_0 \quad (3.17b)$$

$$\alpha \mu_- + \mathcal{L}_f^\dagger \mu_- = \mu_0 \quad (3.17c)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (3.17d)$$

$$\mu_0, \mu_+, \mu_- \in \mathcal{M}_+(X). \quad (3.17e)$$

The supremum in (3.17a) will be attained assuming that X is compact. The bound satisfies $p^* \geq P^*$ as compared to (3.13). Constraints (3.17b) and (3.17c) enforces forward and backwards invariance respectively of trajectories starting from μ_0 . The discount factor α is equivalent to $1/\lambda$, where λ is the scalar value in [39]. The work in [39] only enforces forward invariance, while [37] and equation (3.17a) also enforce backwards invariance.

3.5.2.2 Function Program

The Lagrangian of (3.17a) with dual variables $v_+(x)$, $v_-(x)$ on constraints (3.17b) and (3.17c) is

$$L = \langle p(x), \mu_0 \rangle + \gamma(1 - \langle 1, \mu_0 \rangle) + \langle v_+(x), -\alpha\mu_+ + \mathcal{L}_f^\dagger\mu_+ + \mu_0 \rangle + \langle v_-(x), -\alpha\mu_- - \mathcal{L}_f^\dagger\mu_- + \mu_0 \rangle. \quad (3.18)$$

The function program formed by infimizing L is

$$d^* = \inf_{\gamma \in \mathbb{R}} \gamma \quad (3.19a)$$

$$\gamma \geq v_+(x) + v_-(x) + p(x) \quad \forall x \in X \quad (3.19b)$$

$$\alpha v_+(x) - \mathcal{L}_f v_+(x) \geq 0 \quad \forall x \in X \quad (3.19c)$$

$$\alpha v_-(x) + \mathcal{L}_f v_-(x) \geq 0 \quad \forall x \in X \quad (3.19d)$$

$$v_+(x), v_-(x) \in C^1(X). \quad (3.19e)$$

The mass of μ_0 is constrained to 1 by constraint (3.17d). The occupation measures μ_+ , μ_- each have mass $1/\alpha < \infty$ by definition (3.15). Strong duality with $p^* = d^*$ holds between problems (3.17a) and (3.5) by Theorem 2.6 of [40] because X is compact, all measures are bounded, and the affine maps are continuous. The sublevel set $\{x \mid v_+(x) + v_-(x) \geq 0\}$ is invariant and contains \mathcal{A} .

3.5.2.3 Recovery

The recovery algorithm for global attractors may proceed by a minor modification of Algorithm 1. Let $\mathbb{M}_d[\mathbf{y}^0]$ be the moment matrix associated with the measure μ^0 in the degree- d LMI relaxation of (3.17a). Candidate values for extremizing points of $p(x)$ with $x \in \mathcal{A}$ are the atoms of $\mathbb{M}_d[\mathbf{y}^0]$. An atom x_0 that satisfies $p(x_0) \approx P^*$, such that the trajectory of f starting from x_0 return to an ϵ -neighborhood of x_0 an infinite number of times (x_0 is in its own α and ω limit set [41]), is an approximate optimum of the global attractor peak estimation problem.

3.5.2.4 Global Attractor Examples

The Van-der-Pol oscillator pictured in Figure 3.4 has dynamics

$$\dot{x} = \begin{bmatrix} 2x_2 \\ -0.8x_1 - 10(x_1^2 - 0.21)x_2 \end{bmatrix}. \quad (3.20)$$

CHAPTER 3. PEAK ESTIMATION AND RECOVERY

An order-6 LMI relaxation to program (3.17a) maximizing x_1^2 over the attractive set of the Van-der-Pol oscillator yields a bound of $p_6^* = 0.8585$. The point approximately maximizing $p(x)$ is $x_0 = \pm[0.9266, 0.0030]$.

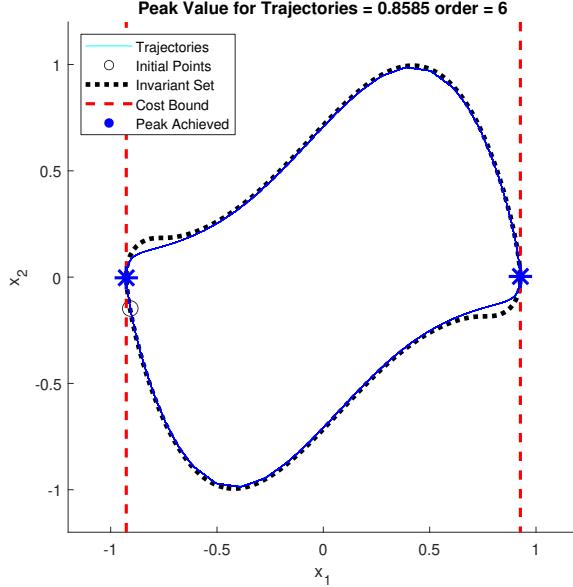


Figure 3.4: Maximize x_1^2 on Van-der-Pol system in (3.20)

The Lorenz attractor is an ODE described by parameters (ρ, σ, β) with behaviour

$$\dot{x} = \begin{bmatrix} \sigma(x_2 - x_1) \\ x_1(\rho - x_3) - x_2 \\ x_1x_2 - \beta x_3 \end{bmatrix}. \quad (3.21)$$

The parameters that Lorenz used (now standard) are $\rho = 28, \sigma = 10, \beta = 8/3$. A scaling of coordinates from [37] of $\tilde{x}_1 = x_1/25, \tilde{x}_2 = x_2/30, \tilde{x}_3 = x_3/50$ may be performed to ensure that the attractor remains within the box $[-1, 1]^3$. An order-6 of maximizing $p(x) = x_1 = 25\tilde{x}_1$ on the attractor of the Lorenz system is visualized in Figure 3.5. The discovered bound is $p_6^* = 18.646$, and the extracted peak point from the nearly rank-1 initial moment matrix (largest two eigenvalues are 2.640, $1.609e-4$) is $x_0^* \approx [0.7458, 0.6309, 0.8280]$. This point is close to the attractor but is not on it, as backwards trajectories traced by MATLAB's `ode45` starting from x_0^* explode around time $t = 1.48$. The global attractor recovery algorithm still yielded a peak point that is likely in the same vicinity as the true bound.

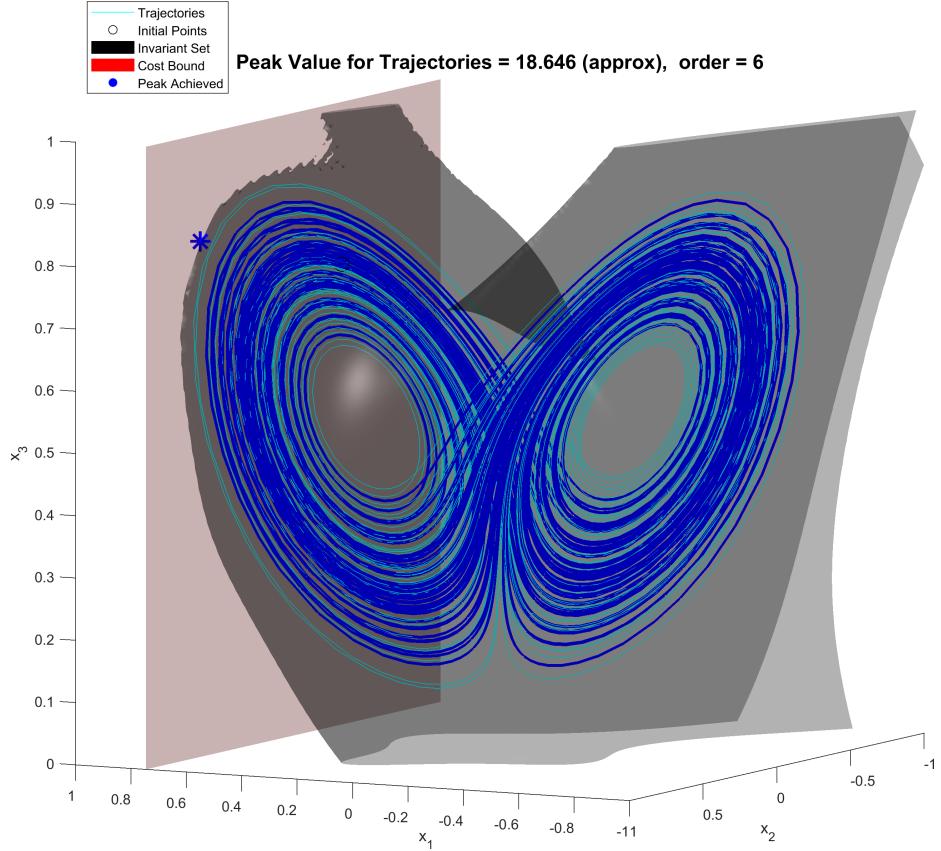


Figure 3.5: Maximize $25\tilde{x}_1$ on the scaled Lorenz system in (3.21)

3.6 Conclusion

This chapter reviewed the approximation of peak estimation programs by LMIs of occupation measure LPs. It also presented the rank-based recovery Algorithm 1 that attempts extraction of optimal trajectories peak estimation problems. Algorithm 1 is applied to problems in ODE peak estimation as well as in finding extrema of global attractors. Future work may involve quantifying the optimality gap between recovered trajectories and the true trajectories, both in terms of difference in peak values, and in normed distance between initial/peak points.

Chapter 4

Peak Estimation under Uncertainty

4.1 Introduction

Peak estimation under uncertainty aims to bound extreme values of a state function subject to an adversarial noise process. Examples include finding the maximum height of an aircraft subject to wind, the maximum voltage in a transmission line subject to thermal noise, and the maximum speed of a motor subject to impedance within a tolerance. A system with finite-dimensional state $x \in \mathbb{R}^{N_x}$ evolves under Ordinary Differential Equation (ODE) dynamics defined by a locally Lipschitz vector field f perturbed by uncertainty over the time-range $t \in [0, T]$. The time-independent uncertainty $\theta \in \Theta \subset \mathbb{R}^{N_\theta}$ is fixed (such as the unknown mass of a system component within tolerance), while the time-dependent uncertainty $w(t)$ may change arbitrarily in time within the region $W \subset \mathbb{R}^{N_w}$. Let $x(t | x_0, \theta, w(t))$ denote a trajectory in time starting from an initial point x_0 subject to uncertainties $(\theta, w(t))$. The uncertain peak estimation problem with variables $(t, x_0, \theta, w(t))$ may be posed as

$$\begin{aligned} P^* &= \sup_{t \in [0, T], x_0 \in X_0, \theta \in \Theta, w(t)} p(x(t | x_0, \theta, w(t))) \\ \dot{x}(t) &= f(t, x(t), \theta, w(t)), \quad w(t) \in W \end{aligned} \tag{4.1}$$

This chapter produces an infinite-dimensional LP in occupation measures to upper bound the quantity P^* from (4.1).

Occupation measure-based bounds for uncertain peak estimation may be developed by adapting methods from optimal control [7, 16]. Time-dependent uncertainty is an instance of an adversarial optimal control which aims to maximize the state function. Time-independent parameter uncertainty may be incorporated by adding states, and switched systems can be analyzed by splitting

the occupation measure [42]. The true peak cost P^* is upper bounded with an infinite-dimensional LP in occupation measures. The infinite LP is then truncated into a sequence of LMIs by the moment-SOS hierarchy [27].

This chapter has the following structure: Section 4.2 reviews uncertainty models. Section 4.3 introduces a unified uncertain peak estimation. Section 4.4 extends uncertain peak estimation to discrete systems. The chapter is concluded in Section 4.5. This chapter's content is from [43] (IEEE CDC Outstanding Student Paper Award), and was coauthored by Didier Henrion, Milan Korda, and Mario Sznaier.

4.2 Uncertainty Models

This section summarizes techniques for incorporating uncertainty into occupation-measure based frameworks, and briefly notes their application to peak estimation. The methods mentioned here arose from optimal control and the approximation of reachability sets. The two basic types of uncertainty are time-independent ($\theta \in \Theta$) and time-dependent ($w \in W$). It is assumed that Θ and W are compact basic semialgebraic sets, just like X and X_0 .

4.2.1 Time-Independent Uncertainty

Time-independent uncertainty θ_ℓ for $\ell = 1..N_\theta$ may take values in a set $\Theta \subseteq \mathbb{R}^{N_\theta}$, and typically arises in systems with parameter tolerances. The time-independent θ may start at any value in $\Theta \subset \mathbb{R}^{N_\theta}$, and is then constant along trajectories. By the methods in [16, 42], the state space may be extended into $X \times \Theta$ by adding new states θ with constant dynamics $\dot{\theta}_\ell = \mathcal{L}_f \theta_\ell = 0$ for each $\ell = 1..N_\theta$.

4.2.2 Time-Dependent Uncertainty

Systems with time-dependent uncertainty may have the noise process $w(t)$ change arbitrarily quickly in W over time t . Such bounded time-varying noise may be found in driving or piloting tasks with changing winds. The disturbance $w(t)$ is a Borel measurable function of time, rather than the Itô-type stochastic process considered in [44] (and discussed later in Chapter 10). For an input $w(t) \in W$ and a subset $D \subseteq W$, the disturbance-occupation measure $\mu^w(A \times B \times D)$ is:

$$\int_{[0,T] \times X_0} I_{A \times B \times D}((t, x(t), w(t)) \mid x_0) dt d\mu_0(x_0). \quad (4.2)$$

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

The disturbance $w(t)$ may be relaxed into a distribution $\omega(w \mid x, t)$, which is known as a Young Measure [45, 7]. The disturbance-occupation measure μ^w can be disentangled into $d\mu^w(t, x, w) = dt d\xi(x \mid t) d\omega(w \mid x, t)$ for conditional distributions ξ, ω . Liouville's equation with a relaxed disturbance $\omega(w \mid x, t)$ influencing dynamics $f(t, x, w)$ for all $v(t, x) \in C^1([0, T] \times X)$ is

$$\langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle \mathcal{L}_f v(t, x), \mu^w \rangle. \quad (4.3a)$$

Equivalent expressions are formed by rearranging operators

$$\langle v, \mu_p \rangle = \langle v, \delta_0 \otimes \mu_0 \rangle + \langle \mathcal{L}_f v, \mu^w \rangle \quad \forall v \quad (4.3b)$$

$$\langle v, \mu_p \rangle = \langle v, \delta_0 \otimes \mu_0 \rangle + \langle v, \mathcal{L}_f^\dagger \mu^w \rangle \quad \forall v \quad (4.3c)$$

$$\langle v, \mu_p \rangle = \langle v, \delta_0 \otimes \mu_0 + \pi_\#^{tx} \mathcal{L}_f^\dagger \mu^w \rangle \quad \forall v. \quad (4.3d)$$

The measures of the two summands on the right-hand side of (4.3c) reside in different spaces, as $\delta_0 \otimes \mu_0 \in \mathcal{M}_+([0, T] \times X)$, while $\mathcal{L}_f^\dagger \mu^w \in \mathcal{M}_+([0, T] \times X \times W)$. The (t, x) -marginalization $\pi_\#^{tx} \mathcal{L}_f^\dagger \mu^w \in \mathcal{M}_+([0, T] \times X)$ allows the measures to be added together inside the duality pairing in (4.3d). The duality pairings $\langle v(t, x), \mathcal{L}_f^\dagger \mu^w \rangle$ and $\langle v(t, x), \pi_\#^{tx} \mathcal{L}_f^\dagger \mu^w \rangle$ are equal for all $v \in C^1([0, T] \times X)$ because $v(t, x)$ is not a function of w . The weak disturbed Liouville's Equation is derived from (4.3d) by treating $\forall v(t, x) \in C^1([0, T] \times X)$ as the implicit expression

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx} \mathcal{L}_f^\dagger \mu^w. \quad (4.4)$$

Time-varying disturbances may be incorporated into peak estimation by letting $\mu \in \mathcal{M}_+([0, T] \times X \times W)$ be a disturbance-occupation measure of the form in (4.2) obeying a disturbed Liouville equation (4.4). The support sets of the measures $\mu_0 \in \mathcal{M}_+(X_0)$, $\mu_p \in \mathcal{M}_+([0, T] \times X)$ are unchanged when time-dependent uncertainty is added.

4.2.3 Switching Uncertainty

An approach for analyzing switched systems with occupation measures is presented in [42]. Let $\{X^k\}_{k=1}^{N_s}$ be a closed cover of X with N_s switching modes. The sets X^k are not necessarily disjoint, and together satisfy $\cup_k X^k = X$ (definition of closed cover). Each region X^k has dynamics $\dot{x} = f_k(t, x)$ for some locally Lipschitz vector field f_k . The closed cover formalism generalizes partitions of X (deterministic dynamics) and arbitrary switching where $X^k = X \ \forall k$ (polytopic uncertainty). Polytopic uncertainty is a model with dynamics $f(t, x, k) = \sum_k w_k f_k(t, x)$ where the disturbance $w_k \in \mathbb{R}_+^{N_s}$ satisfies $\sum_k w_k = 1$. Trajectories from a switching system are equipped

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

with a right-continuous function $S : [0, T] \rightarrow 1..N_s$ yielding the resident subsystem at time t^- . Such a trajectory under switching may be written as $x(t | x_0, S(t))$. The switched measure program introduces an occupation measure $\mu_k \in \mathcal{M}_+([0, T] \times X^k)$ for each subsystem f_k :

$$\mu = \sum_k \mu_k \quad \mathcal{L}^\dagger \mu = \sum_k \mathcal{L}_k^\dagger \mu_k. \quad (4.5)$$

A valid auxiliary function $v(t, x)$ from (3.5c) must decrease along all subsystems [46, 47]. Problem (3.5) may be modified for switching by enlarging Constraint (3.5c) to

$$\mathcal{L}_{f_k} v(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X_k, \quad k = 1..N_s. \quad (4.6)$$

Remark 4.2.1. *The closed cover switching formalism may be expanded into a system with general time-dependent uncertainty if desired. The switching basic semialgebraic sets may be described as $X^k = \{x \mid g_{ki}(x) \geq 0 \text{ } i = 1..N_c^k\}$ for N_c^k polynomial constraints each. A linear expression of time-dependent uncertain dynamics is $\dot{x}(t) = \sum_{k=1}^{N_s} w_k(t) f_k(t, x(t))$ for processes $w(t) \in \mathbb{R}_+^{N_s}$ satisfying $\sum_k w_k(t) = 1$ for all $t \in [0, T]$. Additional constraints must be imposed to enforce that the process $w_k(t)$ is zero whenever $x(t) \notin X^k$. These constraints may be realized as $\{w_k g_{ki}(x) \geq 0, \forall i = 1..N_c^k, \forall k = 1..N_s\}$, given that $x(t) \notin X^k$ if any of the $g_{ki}(x)$ are negative.*

4.3 Continuous-Time Uncertain Peak Estimation

This section combines the uncertainty formulations from Section 4.2 to form a pair of primal-dual infinite-dimensional LPs. The variables $\theta \in \Theta, w \in W$ will respectively denote time-independent and time-dependent uncertainties of sizes N_θ, N_w . The dynamics f have N_s switching subsystems $f_k(t, x, \theta, w)$ which are valid in regions $X_k \subseteq X$.

4.3.1 Continuous-Time Measure Program

A combined uncertain peak estimation measure program is detailed in Program (4.7) with indices $k = 1..N_s$ for the switching subsystems

$$p^* = \sup \quad \langle p(x), \mu_p \rangle \quad (4.7a)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \sum_k \pi_\#^{tx\theta} \mathcal{L}_{f_k}^\dagger \mu_k \quad (4.7b)$$

$$\mu_0(X_0) = 1 \quad (4.7c)$$

$$\mu_k \in \mathcal{M}_+([0, T] \times X_k \times \Theta \times W) \quad \forall k = 1..N_s \quad (4.7d)$$

$$\mu_p \in \mathcal{M}_+([0, T] \times X \times \Theta) \quad (4.7e)$$

$$\mu_0 \in \mathcal{M}_+(X_0 \times \Theta). \quad (4.7f)$$

Theorem 4.3.1. *The solution p^* to program (4.7) will yield an upper bound to P^* in (4.1).*

Proof. First assume $N_s = 1$ with $X^1 = X$, so there is only one switching domain. An optimal achievement of (4.1) reaching the peak value of P^* may be characterized by the tuple $(x_0^*, t^*, x_p^*, \theta^*, w^*(t))$. The peak value $p(x_p^*) = P^*$ is achieved by following the trajectory $x(t \mid x_0^*, \theta^*, w^*(t))$ until time $t = t^*$. Measures (μ_0, μ_p, μ) may be defined from this optimal tuple such that the measures satisfy constraints (4.7b)-(4.7f). The initial measure and peak measure may be set to $\mu_0 = \delta_{x=x_0^*}$ and $\mu_p = \delta_{t=t^*} \otimes \delta_{x=x_p^*} \otimes \delta_{\theta=\theta^*}$ based on the optimal tuple. The measure $\mu \in \mathcal{M}_+([0, T] \times X \times \Theta \times W)$ is the occupation measure of $t \mapsto (t, x(t \mid x_0^*, \theta^*, w^*(t)), \theta^*, w^*(t))$ in the times $[0, t^*]$. The measures (μ_0, μ_p, μ) satisfy constraints (4.7b)-(4.7f), so $p^* \geq P^*$ when $N_s = 1$.

Optimal trajectories arising from a system with $N_s > 1$ may be described in a tuple as $(x_0^*, t^*, x_p^*, \theta^*, w^*(t), S^*(t))$, where $S^*(t)$ is the sequence of switches undergone between times $t \in [0, t^*]$. The measures μ_0 and μ_p may remain the same as in the non-switched case. Switching occupation measures μ_k may be set to the unique occupation measure supported on the graph $(t, x(t \mid x_0^*, \theta^*, w^*(t)), \theta^*, w^*(t))$ between times $t \in [0, t^*]$ when $S(t) = k$. These occupation measures satisfy constraints (4.7b) and (4.7d), proving that there exists a feasible solution to (4.7b)-(4.7f) for the case of switching with objective P^* . \square

Theorem 4.3.2. *Assumptions that the sets $X_0, \forall k : X_k, \Theta, W, [0, T]$ are compact, that p is continuous, and that each f_k is Lipschitz within $[0, T] \times X_k \times \Theta \times W$. Further assume that the image $f_\ell(t, x, \theta, W)$ is convex for each fixed $(t, x, \theta) \in [0, T] \times X \times \Theta$ and $\ell \in 1..N_s$. Then program (4.7) has the same optimal value as (4.1) with $p^* = P^*$.*

Proof. Program (4.1) is an instance of an Optimal Control Problem (OCP) with zero running cost and free terminal time. Theorem 2.1 of [7] ensures that the measure LP (4.7) will have the same optimal value as the OCP (4.1) under the provided compactness, continuity, regularity, and dynamics-convexity assumptions. \square

4.3.2 Continuous-Time Function Program

Dual variables $v(t, x, \theta) \in C^1([0, T] \times X \times \Theta)$ and $\gamma \in \mathbb{R}$ can be defined to find the Lagrangian of (4.7)

$$\begin{aligned}\mathcal{L} = & \langle p(x), \mu_p \rangle + \langle v(t, x, \theta), \delta_0 \otimes \mu_0 + \sum_k \pi_{\#}^{tx\theta} \mathcal{L}_{f_k}^\dagger \mu_k \rangle \\ & + \langle v(t, x, \theta), -\mu_p \rangle + \gamma(1 - \langle 1, \mu_0 \rangle).\end{aligned}$$

The resulting dual program in (v, γ) is

$$d^* = \inf_{\gamma \in \mathbb{R}} \quad \gamma \tag{4.8a}$$

$$\forall (x, \theta) \in X_0 \times \Theta :$$

$$\gamma \geq v(0, x, \theta) \tag{4.8b}$$

$$\forall (t, x, \theta, w) \in [0, T] \times X_k \times \Theta \times W, \quad \forall k :$$

$$\mathcal{L}_{f_k} v(t, x, \theta) \leq 0 \tag{4.8c}$$

$$\forall (t, x, \theta) \in [0, T] \times X \times \Theta :$$

$$v(t, x, \theta) \geq p(x) \tag{4.8d}$$

$$v(t, x, \theta) \in C^1([0, T] \times X \times \Theta). \tag{4.8e}$$

Theorem 4.3.3. *There is no duality gap between (4.7) and (4.8) when the set $[0, T] \times X \times \Theta \times W$ is compact.*

Proof. Necessary and sufficient conditions for there to be no duality gap between measure and function programs are if the objective p^* is bounded and if the affine map is closed in the weak-* topology (Theorem 2.6 in [40]). The function $p(x)$ is bounded over the compact set X and μ_p is a probability distribution, so the objective $\langle p(x), \mu_p \rangle$ is therefore bounded. The image of the affine map $(\mu_0, \mu_p, \mu_k) \rightarrow (\delta_0 \otimes \mu_0 + \sum_k \pi_{\#}^{tx\theta} \mathcal{L}_{f_k}^\dagger \mu_k - \mu_p, \mu_0)$ induced by constraints (4.7b)-(4.7c) is closed in the weak-* topology [8]. Strong duality therefore holds by closure and boundedness of measures. \square

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

The measure μ_0 has $N_x + N_\theta$ variables, and μ_p has $1 + N_x + N_\theta$ variables. The N_s occupation measures μ_k each have $1 + N_x + N_\theta + N_w$ variables. If the switching structure was not taken into account by the methods of section 4.2.3, there would be a single occupation measure μ with $1 + N_x + N_\theta + N_w + N_s$ variables. The affine uncertainty structure breaks up the large μ (in terms of the number of variables) into N_s smaller measures (μ_k).

4.3.3 Continuous-Time LMI Relaxation

The compact (Archimedean) basic semialgebraic sets in the uncertain peak estimation setting are

$$X = \{x \mid g_i(x) \geq 0 \quad | \quad i = 1..N_c\} \quad (4.9a)$$

$$X_0 = \{x \mid g_{0i}(x) \geq 0 \quad | \quad i = 1..N_c^0\} \quad (4.9b)$$

$$X^k = \{x \mid g_{ki}(x) \geq 0 \quad | \quad i = 1..N_c^k\} \quad (4.9c)$$

$$\Theta = \{\theta \mid g_{\theta i}(\theta) \geq 0 \quad | \quad i = 1..N_c^\theta\} \quad (4.9d)$$

$$W = \{w \mid g_{wi}(w) \geq 0 \quad | \quad i = 1..N_c^w\}. \quad (4.9e)$$

The degree of $g_i(x)$ is d_i , and other degrees d_{0i} , $d_{\theta i}$, d_{wi} , d_{ki} are defined on corresponding polynomials. Monomials forming moments may be indexed as $x^\alpha t^\beta \theta^\gamma w^\eta$ for multi-indices $\alpha \in \mathbb{N}^{N_x}$, $\beta \in \mathbb{N}$, $\gamma \in \mathbb{N}^{N_\theta}$, $\eta \in \mathbb{N}^{N_w}$. Define $\mathbf{m}^0 = \{\mathbf{m}_{\alpha\gamma}^0\}$, $\mathbf{m}^p = \{\mathbf{m}_{\alpha\beta\gamma}^p\}$ as the moment sequences for measures μ_0 and μ_p . The moment sequence for the occupation measure μ^k is $\mathbf{m}^k = \{\mathbf{m}_{\alpha\beta\gamma\eta}^k\}$ for each switching subsystem k . The Liouville equation (4.7b) with test function $v(t, x, \theta) = x^\alpha t^\beta \theta^\gamma$ has the form

$$0 = \langle x^\alpha t^\beta \theta^\gamma, \delta_0 \otimes \mu_0 \rangle - \langle x^\alpha t^\beta \theta^\gamma, \mu_p \rangle + \sum_k \langle \mathcal{L}_{f_k(t, x, \theta, w)}(x^\alpha t^\beta \theta^\gamma), \mu_k \rangle. \quad (4.10)$$

Define the operator $\text{Liou}_{\alpha\beta\gamma}(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^k)$ as the linear relation between the moment sequences induced by (4.10) assuming that each f_k is a polynomial vector field. Given a degree d , define the degrees d'_k as $d + \lceil \deg(f_k)/2 \rceil - 1$ for each k . The degree- d LMI relaxation of the uncertain peak

estimation problem in (4.7) resulting in an upper bound $p_d^* \geq P^*$ is

$$p_d^* = \max \quad \sum_{\alpha} p_{\alpha} y_{\alpha 00}^p \quad (4.11a)$$

$$\text{Liou}_{\alpha\beta\gamma}(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^k) = 0 \quad \text{by (4.10)} \quad \forall |\alpha| + |\beta| + |\gamma| \leq 2d \quad (4.11b)$$

$$\mathbf{m}_0^0 = 1 \quad (4.11c)$$

$$\mathbb{M}_d((X_0 \times \Theta)\mathbf{m}^0), \mathbb{M}_d(([0, T] \times X \times \Theta)\mathbf{m}^p), \succeq 0 \quad (4.11d)$$

$$\forall k : \mathbb{M}_{d'_k}(([0, T] \times X_k \times \Theta \times W)\mathbf{m}^k) \succeq 0. \quad (4.11e)$$

Constraints (4.11d)-(4.11e) are moment and localizing matrix PSD constraints ensuring that there exist representing measures to the moment sequences $(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^k)$ supported on the appropriate spaces.

In Program (4.7), μ_0 and μ_p each have mass 1, and the mass of $\sum_k \mu_k \leq T$ by Liouville's equation. Compactness of $[0, T] \times X \times \Theta \times W$ therefore assures that all measures are bounded. Given this boundedness, sequence $\{p_d^*\}$ will converge to p^* monotonically from above as $d \rightarrow \infty$ if all sets in (4.9) are Archimedean [27].

Algorithm 1 may be used to attempt localization of peak points (t_p, x_p, θ) and initial points $(0, x_0, \theta)$. Algorithm 1 can recover near-optimal values for time-independent uncertainty $\theta \in \Theta$ if the rank condition holds, but it is unable to determine the optimal time-dependent noise process $w(t) \in W$, nor the optimal switching sequence $S(t)$. The work in [31] that estimates support of a measure from its moments may be used to approximate the noise processes $w(t)$ and $S(t)$ from the w -marginals $\pi_{\#}^w \mu^k$ (appropriate moments of which are contained in $\mathbb{M}_{d'_k}(\mathbf{m}^k)$).

4.3.4 Continuous-Time Uncertain Examples

Code is available at github.com/jarmill/peak, and is written in MATLAB R2020a using Gloptipoly3 [30], YALMIP [48], and Mosek 9.2 [49] to formulate and solve LMIs. Demonstrations are available in the folder `peak/experiments_uncertain`, and were run on an Intel i9 CPU at 2.30 GHz with 64.0 GB of RAM.

4.3.4.1 Inner Tube

Dynamics based on Example 1 of [50] (adding w) are

$$\dot{x}(t) = \begin{bmatrix} -0.5x_1 - (0.5 + w(t))x_2 + 0.5 \\ -0.5x_2 + 1 + \theta \end{bmatrix}. \quad (4.12)$$

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

Figure 4.1 illustrates maximization of $p(x) = x_1$ starting in $X_0 = \{x \mid (x_1 + 1)^2 + (x_2 + 1)^2 \leq 0.25\}$ for time $t \in [0, 10]$. The admissible disturbances $w(t)$ are in $W = [-0.2, 0.2]$. Fig. 4.1a has $\Theta = 0$, while Fig. 4.1b has $\Theta = [-0.5, 0.5]$ for the time-independent uncertainty $\theta \in \Theta$. In each figure, the black circles are initial conditions from the boundary of X_0 , the blue curves are sampled trajectories, and the red plane are level sets for upper bounds of x_1 along trajectories. At the order $r = 4$ LMI relaxation, Fig. 4.1a yields a bound of $P^* \leq 0.4925$, while Fig. 4.1b with θ results in $P^* \leq 0.7680$. The black surface containing all trajectories in Fig. 4.1a is the level set $\{(t, x) \mid v(t, x) = 0.4925\}$.

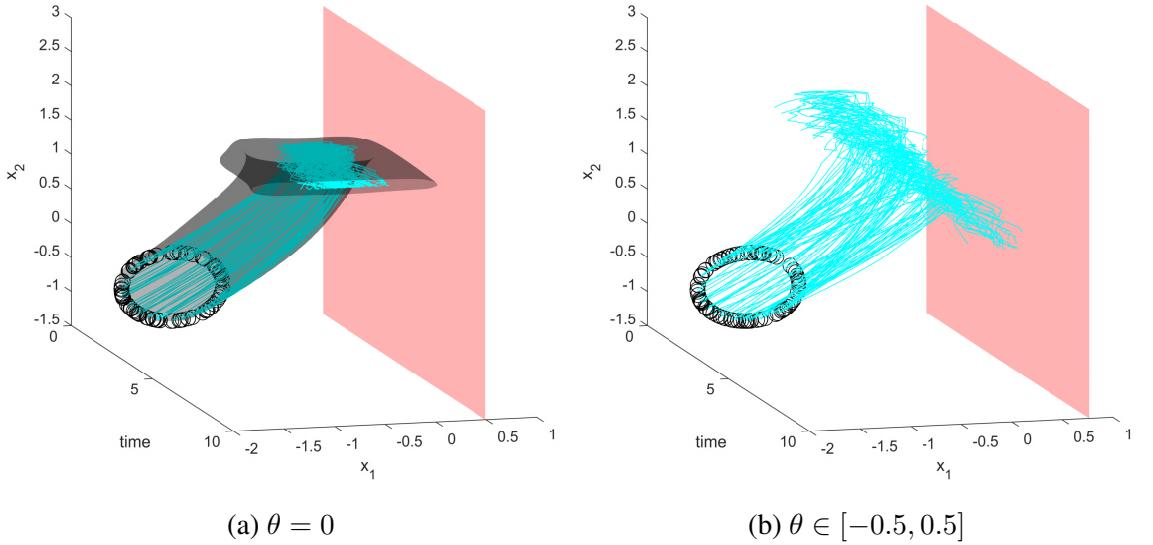


Figure 4.1: Maximize x_1 at order 4 with $w(t) \in W$

4.3.4.2 Three-Wave

The reduced three-wave model is a nonlinear model for the interaction of three quasisynchronous waves in a plasma [51]. These dynamics with parameters (A, B, G) are

$$\begin{aligned} \dot{x}_1 &= Ax_1 + Bx_2 + x_3 - 2x_2^2 \\ \dot{x}_2 &= -Bx_1 + Ax_2 + 2x_1x_2 \\ \dot{x}_3 &= -Gx_3 - 2x_1x_2. \end{aligned} \tag{4.13}$$

This example aims to maximize x_2 on the three-wave system starting in $X_0 = \{x \mid (x_1 + 1)^2 + (x_2 + 1)^2 + (x_3 + 1)^2 \leq 0.16\}$. Order 3 LMI relaxations are used to upper bound x_2 over the

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

region of interest $X = [-4, 3] \times [0.5, 3.6] \times [0, 4]$ and times $t \in [0, 5]$. The bound $P^* \leq 2.6108$ is produced with parameter values $A = 1$, $B = 0.5$, $G = 2$ (no uncertainty), as illustrated in Fig. 4.2a. Fig. 4.2b adds uncertainty by letting $A \in [-0.5, 1.5]$ and $B \in [0.25, 0.75]$ vary arbitrarily with time, and G now possesses parametric uncertainty in $[1.9, 2.1]$. Uncertainty in A, B are realized by switching between 4 subsystems of (4.13) with $(A, B) \in \{0.5, 1.5\} \times \{0.25, 0.75\}$. Uncertainty in G is implemented as $G = 2 + \theta$ where $\theta \in [-0.1, 0.1]$. The order-3 bound under uncertainty in Fig. 4.2b is $P^* \leq 3.296$.

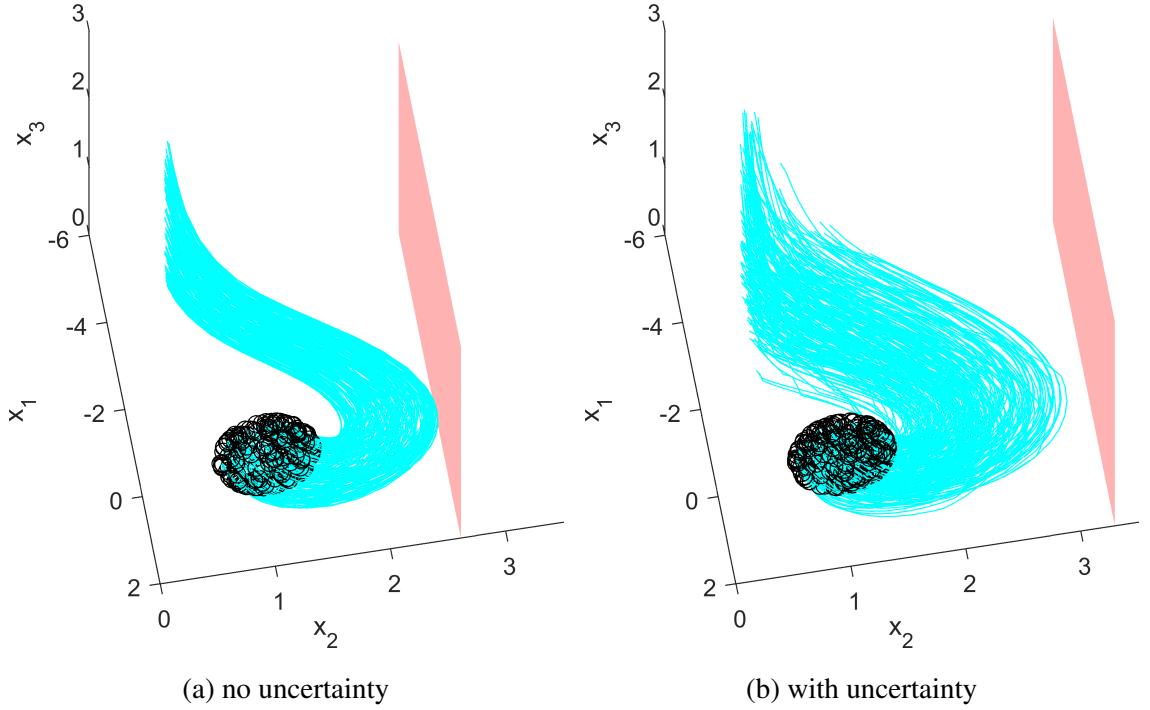


Figure 4.2: Maximize x_2 on three-wave system (4.13)

4.3.4.3 Spacecraft Attitude

Section 4.1 of [42] introduces a 1DOF attitude controller for validation of a space launcher system. These linearized dynamics corresponding to a double-integrator $I\ddot{\phi} = u$ and states $x = [\phi, \dot{\phi}]$. The input $u = \text{sat}_L(Kx)$ is a state feedback controller $Kx = 1000(2.475\phi + 19.8\dot{\phi})$ that saturates at levels $\pm L = \pm 380$. The subsystems are linear operation $|Kx| \leq L$, positive saturation $Kx \geq L$, and negative saturation $Kx \leq -L$ (deterministic switching). These valid regions X^k are separated in Fig. 4.3 by thin dotted diagonal lines. Maximizing $p(x) = |\phi|$ (implemented as ϕ^2) is shown in Fig.

4.3a. With $|\phi_0| \leq 15^\circ$ and $|\dot{\phi}_0| \leq 3^\circ/\text{sec}$, a degree-5 approximation finds a time-independent upper bound of $|\phi_*| = 20.69^\circ$. The blue curve is the near-optimal trajectory, starting at the blue circle and extremizing $p(x)$ at the blue star. The nominal moment of inertia in Fig. 4.3a is $I = 27,500 \text{ kg m}^2$. Time-independent relative uncertainty I may be introduced by replacing I with $I/(1 + \theta)$, where $\theta \in [-0.5, 0.5]$. The peak angle is raised to $|\phi_*| = 51.86^\circ$ at $d = 5$ at $I' = 2I$ with this new uncertainty in Fig. 4.3b.

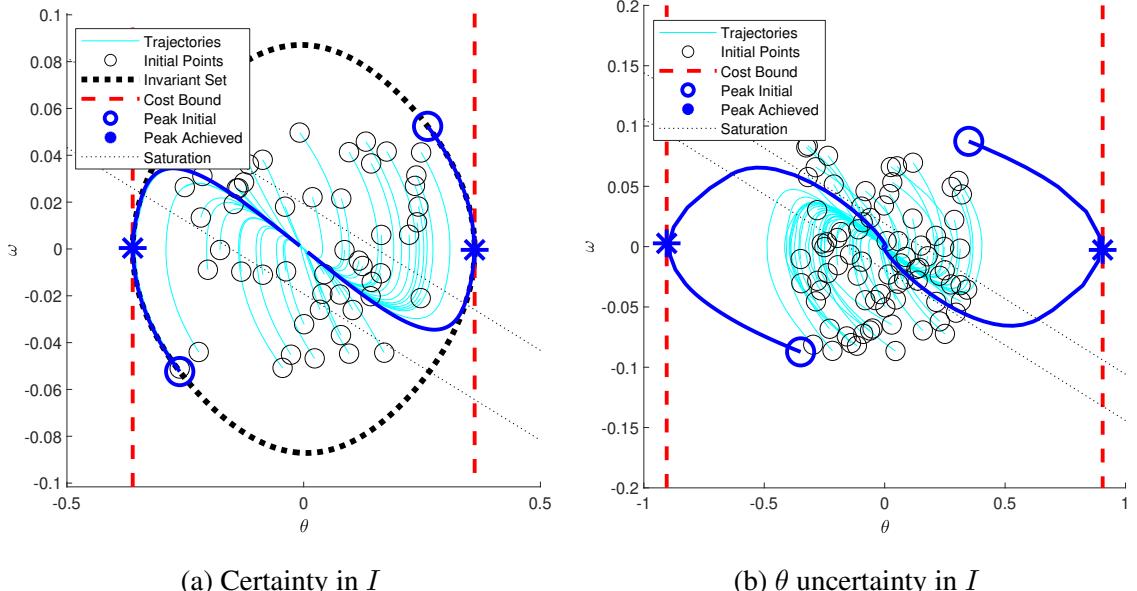


Figure 4.3: Maximum angle for 1DOF attitude controller

4.4 Discrete-Time Uncertain Peak Estimation

Uncertain peak estimation can be extended to discrete systems, including switched discrete-time systems. A discrete-time system from times $t = 0..T$ is considered for dynamics $x_+ = f(x)$ where x_+ is the next state. A trajectory starting at the initial condition $x_0 \in X_0$ is $x_t(x_0)$. The uncertain peak estimation problem for discrete-time systems with uncertainties (θ, w_t) and N_s subsystems with switching sequence S_t is

$$P^* = \sup_{t, x_0 \in X_0, \theta \in \Theta, w_t, S_t} p(x_t(x_0, \theta, w_t, S_t)) \quad (4.14)$$

$$x_+ = f_k(x_t, \theta, w_t) \text{ if } S_t = k$$

$$w_t \in W, S_t \in 1..N_s \quad \forall t \in 0..T.$$

4.4.1 Discrete-Time Measure Background

Just as the Lie derivative $\mathcal{L}_f v$ yields the infinitesimal change in v along continuous trajectories, the quantity $v(f(x)) - v(x)$ is the change in v along a single discrete step. An occupation measure for sets $A \subseteq X$ with initial conditions distributed as $\mu_0 \in \mathcal{M}_+(X_0)$ may be defined for discrete systems as

$$\mu(A) = \int_{X_0} \sum_{t=0}^T I_A(f^t(x_0)) d\mu_0. \quad (4.15)$$

The quantity $\mu(A)$ is the averaged number of time steps that trajectories distributed as μ_0 spend in the region A . For measures $\mu_0 \in \mathcal{M}_+(X_0)$, $\mu_p \in \mathcal{M}_+(X)$, $\mu \in \mathcal{M}_+(X)$, the strong and weak discrete Liouville equations for all v are

$$\langle v(x), \mu_p \rangle = \langle v(x), \mu_0 \rangle + \langle v(f(x)), \mu \rangle - \langle v(x), \mu \rangle, \quad (4.16)$$

$$\mu_p = \mu_0 + f_\# \mu - \mu. \quad (4.17)$$

Time may be optionally included in system dynamics by setting a state $t_+ = t + 1$ and incorporating t into dynamics. The pushforward term in (4.17) would then be $v(t + 1, f(t, x)) - v(t, x)$. Discrete systems with uncertainties (θ, w) have dynamics and Liouville equations according to

$$x_+ = f(x_t, \theta, w_t), \quad \mu_p = \mu_0 + \pi_{\#}^{x\theta}(f_\# \mu - \mu). \quad (4.18)$$

The uncertainty $\theta \in \Theta$ is fixed, and the time-dependent uncertainty has $w_t \in W$ for every time step $t = 0..T$. Switching uncertainty from Section 4.2.3 with subsystems f_k valid over X_k may be realized by defining occupation measures $\mu_k \in \mathcal{M}_+(X_k \times \Theta \times W)$ such that $\mu = \sum_k \mu_k$.

4.4.2 Discrete-Time Measure Program

A measure program may be formulated to upper bound the peak-estimation task on discrete systems. The uncertainties available in this formulation are (θ, w) and switching between dynamics

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

f_k over X_k . The uncertain discrete peak estimation measure problem with variables (μ_0, μ_k, μ_p) is

$$p^* = \sup \quad \langle p(x), \mu_p \rangle \quad (4.19a)$$

$$\mu_p = \mu_0 + \pi_{\#}^{x\theta} (\sum_k (f_k \# \mu_k - \mu_k)) \quad (4.19b)$$

$$\mu_0(X_0) = 1 \quad (4.19c)$$

$$T \geq \sum_k \langle 1, \mu_k \rangle \quad (4.19d)$$

$$\mu_k \in \mathcal{M}_+(X_k \times \Theta \times W) \quad \forall k = 1..N_s \quad (4.19e)$$

$$\mu_p \in \mathcal{M}_+(X \times \Theta) \quad (4.19f)$$

$$\mu_0 \in \mathcal{M}_+(X_0 \times \Theta). \quad (4.19g)$$

Remark 4.4.1. *The composition of pushforwards in (4.19b) acts as*

$$\langle v(x, \theta), \pi_{\#}^{x\theta} f_k \# \mu_k \rangle = \langle v(f_k(x, \theta, w), \theta), \mu_k \rangle \quad (4.20)$$

for all test functions $v(x, \theta) \in C(X \times \Theta)$.

Theorem 4.4.1. *The optimum p^* of (4.19) is an upper bound for P^* from discrete program (4.14).*

Proof. This proof follows the same steps as the proof to theorem 4.3.1. An trajectory achieving a peak value of P^* solving (4.14) may be expressed as a tuple $(t^*, x_0^*, x_p^*, \theta^*, w_t^*, S_t^*)$ with $P^* = p(x_p^*) = p(x_{t^*}(x_0^*, \theta^*, w_t^*))$. Measures may be defined from this tuple to solve problem (4.19). The probability distributions are $\mu_0 = \delta_{x=x_0^*}$ and $\mu_p = \delta_{x=x_p^*} \otimes \delta_{\theta=\theta^*}$. Switching measures μ_k may be chosen as the discrete-time occupation measures of $t \mid (x_t(x_0^*, \theta^*, w_t^*), \theta^*, w_t^*) I(S_t = k)$ for $t \in 0..t^*k$. for all test functions $\tilde{v}_k \in C(X_k \times \Theta \times W)$ and for each $k = 1..N_s$. The measures (μ_0, μ_p, μ_k) are feasible solutions to (4.19b)-(4.19g) with objective value $P^* = p(x_p^*) = \langle p(x), \mu_p \rangle$, so $p^* \geq P^*$ is a valid upper bound to (4.14). \square

Remark 4.4.2. *Constraint (4.19d) is a technique from [52] ensuring that the maximal time in optimization is T and that each μ_k has a bounded mass.*

4.4.3 Discrete-Time Function Program

With dual variables $(v(x, \theta) \in C(X \times \Theta), \gamma \in \mathbb{R})$ and a new dual variable $\alpha \geq 0$, the Lagrangian of (4.19) is

$$\begin{aligned} \mathcal{L} &= \langle p(x), \mu_p \rangle + \langle v(x, \theta), \mu_0 - \mu_p \rangle + \alpha(T - \langle 1, \sum_k \mu_k \rangle) \\ &\quad + \langle v(x, \theta), \pi_{\#}^{x\theta} \sum_k f_k \# \mu_k - \mu_k \rangle + \gamma(1 - \langle 1, \mu_0 \rangle). \end{aligned}$$

CHAPTER 4. PEAK ESTIMATION UNDER UNCERTAINTY

The corresponding dual problem is

$$d^* = \inf_{\gamma \in \mathbb{R}, \alpha \geq 0} \gamma + T\alpha \quad (4.21a)$$

$$\forall (x, \theta) \in X_0 \times \Theta :$$

$$\gamma \geq v(x, \theta) \quad (4.21b)$$

$$\forall (x, \theta, w) \in X_k \times \Theta \times W : \quad \forall k$$

$$v(f_k(x, \theta, w), \theta) - v(x, \theta) \leq \alpha \quad (4.21c)$$

$$\forall (x, \theta) \in X \times \Theta :$$

$$v(x, \theta) \geq p(x) \quad (4.21d)$$

$$v(x, \theta) \in C(X \times \Theta). \quad (4.21e)$$

Theorem 4.4.2. *Strong duality $p^* = d^*$ between holds between (4.19) and (4.21) if $T < \infty$ and $X \times \Theta \times W$ is compact.*

Proof. This is affirmed by a similar process to Theorem 4.3.3. The objective $\langle p, \mu_p \rangle$ is bounded. Additionally, all measures have bounded finite moments given that their masses are bounded and their supports are compact. The image of the affine map in constraints (4.19b)-(4.19c) is closed in the weak-* topology, concluding the conditions for strong duality by Theorem 2.6 of [40]. \square

4.4.4 Discrete LMI

The LMI relaxation of (4.19) can be developed in the same manner as in Section 4.3.3. The sets (X_0, X, W, D) are defined in the same way as in equation (4.9). As there is no t term in discrete systems, monomials forming moments are indexed as $x^\alpha \theta^\gamma w^\eta$. The moment sequences are y^0, y^p , and a y^k for each switching subsystem $k = 1, \dots, N_s$. The Liouville equation (4.19b) with a given test function $v(x, \theta) = x^\alpha \theta^\gamma$ is

$$0 = \langle x^\alpha \theta^\gamma, \delta_0 \otimes \mu_0 \rangle - \langle x^\alpha \theta^\gamma, \mu_p \rangle + \sum_k \langle (f_k(x, \theta, w))^\alpha \theta^\gamma - x^\alpha \theta^\gamma, \mu_k \rangle. \quad (4.22)$$

The operator $\text{Liou}_{\alpha\gamma}(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^k)$ is defined as the relation induced by the discrete

Liouville equation (4.22). The discrete degree- d LMI truncation of (4.19) is

$$p_d^* = \sup \quad \sum_{\alpha} p_{\alpha} y_{\alpha 0}^p \quad (4.23a)$$

$$\text{Liou}_{\alpha\gamma}(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^k) = 0 \quad \text{by (4.22)} \quad \forall |\alpha| + |\gamma| \leq 2d \quad (4.23b)$$

$$\mathbf{m}_0^0 = 1 \quad (4.23c)$$

$$\sum_k \mathbf{m}_0^k \leq T \quad (4.23d)$$

$$\mathbb{M}_d((X_0 \times \Theta)\mathbf{m}^0), \mathbb{M}_d(([0, T] \times X \times \Theta)\mathbf{m}^p) \succeq 0 \quad (4.23e)$$

$$\forall k : \mathbb{M}_d(([0, T] \times X_k \times \Theta \times W)\mathbf{m}^k) \succeq 0. \quad (4.23f)$$

Constraint (4.23d) enforces the time limit constraint on occupation measures (4.19d). The structure of (4.23e)-(4.23f) is similar to (4.11) with the affine, moment matrix and localizing matrix constraints.

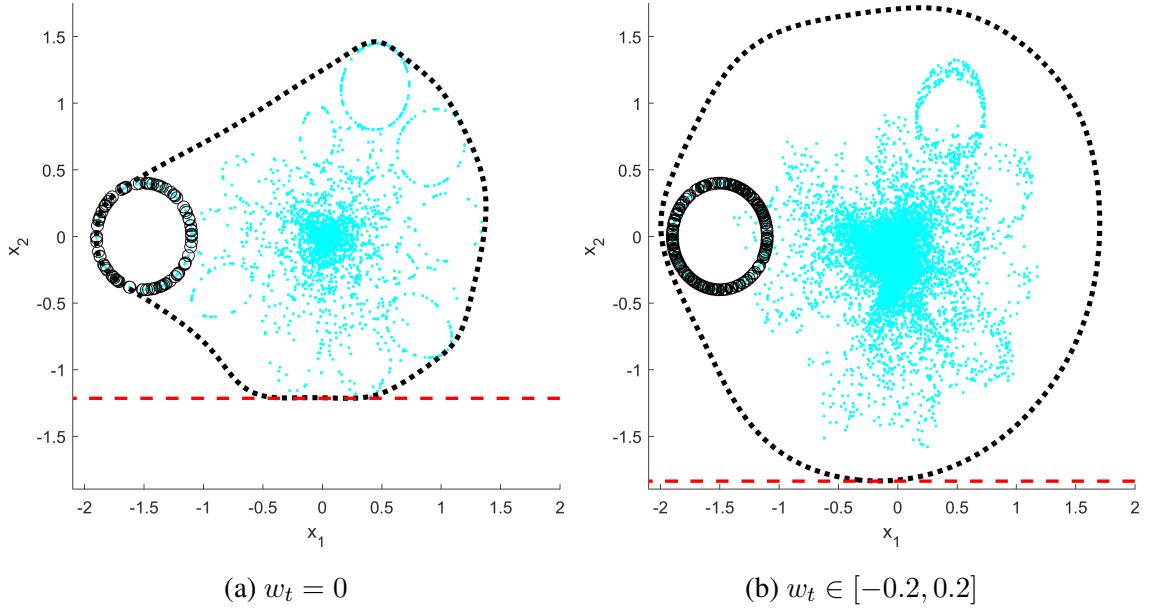
4.4.5 Discrete Example

An example to demonstrate uncertain discrete peak estimation is to minimize x_2 on the following subsystems:

$$f_1(x, w) = \begin{bmatrix} -0.3x_1 + 0.8x_2 + 0.1x_1x_2 \\ -0.75x_1 - 0.3x_2 + w \end{bmatrix} \quad (4.24a)$$

$$f_2(x, w) = \begin{bmatrix} 0.8x_1 + 0.5x_2 - 0.01x_1^2 \\ -0.5x_1 + 0.8x_2 - 0.01x_1x_2 + w \end{bmatrix}. \quad (4.24b)$$

The space under consideration is $X = [-3, 3]^2$, and the time varying uncertainty w_t satisfies $w_t \in [-0.2, 0.2] = W$. The valid regions for subsystems of (4.24) are $X_1 = X$ and $X_2 = X \cap (x_1 \geq 0)$. When $x_1 \geq 0$, the system may switch arbitrarily between dynamics f_1 and f_2 , but when $x_2 < 0$, the system only follows dynamics f_1 . Figure 4.4 visualizes minimizing x_2 starting from the initial set $X_0 = \{x \mid (x_1 + 1.5)^2 + x_2^2 = 0.16\}$ between discrete times $t \in 0, \dots, T$ with $T = 50$. A fourth order LMI relaxation of (4.19a) is solved aiming to maximize $p(x) = -x_2$. With $w = 0$ in Fig. 4.4a the bound is $P^* \leq 1.215$ ($\min x_2 \geq -1.215$), while the time varying w in Fig. 4.4b yields a bound of $P^* \leq 1.837$.


 Figure 4.4: Minimize x_2 on system (4.24)

4.5 Conclusion

The problem of peak estimation with uncertainty may be bounded by the optimal value of an infinite-dimensional LP in occupation measures. Available uncertainty processes discussed in this chapter are time-independent, arbitrarily time-dependent, and switching. This LP is then approximated by the moment-SOS hierarchy and Linear Matrix Inequalities. Time-independent and time-dependent uncertainties are incorporated into this measure framework for continuous and discrete systems. Upcoming chapters of this thesis will focus on cases where the noise has additional decomposable structure (Chapter 6), and will also treat probability-based bounds on peak values under stochastic differential equations (Chapter 10).

Chapter 5

Peak Estimation for Safety Analysis

5.1 Introduction

A trajectory is safe with respect to an unsafe set X_u if no point along trajectories contacts or enters X_u . This chapter presents two methods to certify safety trajectories with respect to X_u : safety margins and distance of closest approach. Safety Margins are a measure of violation for the nonnegativity constraints describing the BSA set X_u . A negative safety margin verifies safety of trajectories starting from X_0 . The safety margin may be computed by extending measure-LP peak estimation techniques to problems with maximin objectives. The distance of closest approach between points along trajectories and X_u will be positive for all safe trajectories, and will be zero for all unsafe trajectories. The task of finding this distance of closest approach will also be referred to as ‘distance estimation’. Distance estimation problems may be solved by using techniques from optimal transport theory, relaxing a distance objective into an expectation of the distance $c(x, y)$ with respect to probability distributions over X and X_u [53, 54, 55].

Prior work on verifying safety of trajectories includes Barrier functions [36, 56] and Density functions [57]. Barrier and Density functions offer binary indications of safety/unsafety; if a Barrier/Density function exists, then all trajectories starting from X_0 are safe. Barrier/Density functions may be non-unique, and the existence of such a function does not yield a measure of closeness to the unsafe set. Safety Margins can vary with constraint reparameterization (e.g., multiplying all defining constraints of X_u by a positive constant) and therefore yield a qualitative certificate of safety. The distance of closest approach P^* is independent of constraint reparameterization and returns quantifiable and geometrically interpretable information about the safety of trajectories.

This chapter is structured as follows: Section 5.2 reviews barrier functions for safety

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

verification. Section 5.3 introduces the concept of safety margins which may be computed through maximin peak estimation. Section 5.4 proposes an infinite-dimensional LP to bound the distance closest approach between points along trajectories and points on the unsafe set. Section 5.5 uses the moment-SOS hierarchy on the distance estimation LP to form a convergent sequence of SDPs. Section 5.6 utilizes correlative sparsity to create SDPs of distance estimation with smaller PSD matrix constraints. Section 5.7 poses distance estimation problems for shapes traveling along trajectories. Section 5.8 presents examples of distance estimation. Section 5.9 details extensions to the distance estimation problem, including uncertainty, polyhedral norm distances, and application of correlative sparsity. Section 5.10 concludes the chapter. Appendix A.1 contains a proof of strong duality for the distance LPs.

The safety margin content of this chapter appeared in [32] and was coauthored by Didier Henrion and Mario Sznaier. The distance estimation work is from [58, 13] and was coauthored by Mario Sznaier.

A persistent example throughout this chapter is verifying safety of the Flow system in Figure 5.1 with respect to the unsafe set $X_u = \{x \mid x_1^2 + (x_2 + 0.7)^2 \leq 0.5^2, \sqrt{2}/2(x_1 + x_2 - 0.7) \leq 0\}$. The set X_u is the red half-circle to the bottom-left of trajectories.

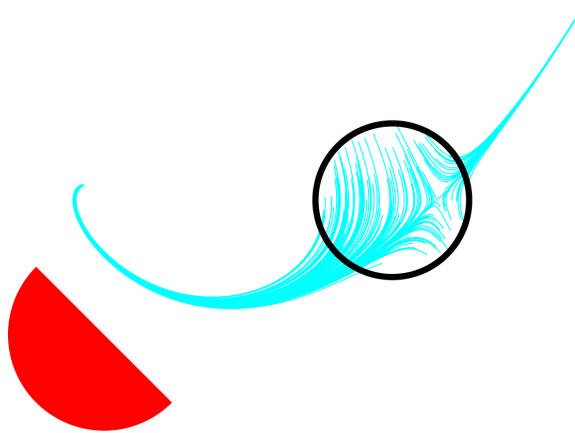


Figure 5.1: Trajectories of Flow system (3.9)

5.2 Barrier Functions

A barrier function $B \in C(X)$ is a continuous function that obeys the following nonnegativity constraints [36, 56]:

$$B(x) \leq 0 \quad \forall x \in X_u \quad (5.1a)$$

$$B(x) > 0 \quad \forall x \in X_0 \quad (5.1b)$$

$$f(t, x) \cdot \nabla_x B(x) \geq 0 \quad \forall x \in X. \quad (5.1c)$$

The barrier function begins positive on X_0 (5.1b) and increases along all trajectories (5.1c). It is therefore not possible for trajectories to visit X_u where the barrier function is negative (5.1a). The existence of a $B(x)$ that solves (5.1) is sufficient to certify safety of trajectories with respect to X_u . Constraint (5.1a) may be relaxed to $B(x) \leq 0 \forall x \in \partial X_u$ with no loss of generality so long as $X_0 \cap X_u = \emptyset$.

Barrier functions are non-unique: for any $r > 0$, the function $rB(x)$ will also be a Barrier function if $B(x)$ is a Barrier function. Barrier functions are strong alternatives to feasibility of the following measure program with variables (μ_0, μ, μ_u) :

$$\mu_u = \mu_0 + \mathcal{L}_f^\dagger \mu \quad \mu_0(X_0) = 1 \quad (5.2a)$$

$$\mu_0 \in \mathcal{M}_+(X_0), \mu_u \in \mathcal{M}_+(X_u), \mu \in \mathcal{M}_+(X). \quad (5.2b)$$

The green curve in Figure 5.2 is the level set $B(x) = 0$ of a degree-6 polynomial barrier function certifying safety of the Flow system, found through a sum-of-squares relaxation of (5.1) with the state constraint $X = [-3, 3]^2$.

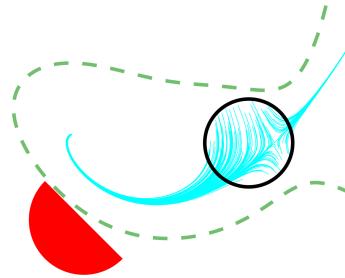


Figure 5.2: Degree-6 Barrier function for Flow system (3.9)

5.3 Safety Margins and Maximin Peak Estimation

5.3.1 Safety Margin Background

Assume that X_u is a BSA set with description $X_u = \{x \mid p_i(x) \geq 0, i = 1..N_u\}$. A point x is in X_u if all $p_i(x)$ are nonnegative. If at least one $p_i(x)$ remains negative for all points along trajectories $x(t \mid x_0), x_0 \in X_0$, then no point starting from X_0 enters X_u and trajectories are safe. The value $p^* = \min_i p_i(x)$ is called the safety margin, and a negative safety margin $p^* < 0$ certifies safety.

5.3.2 Maximin Program

The safety margin estimation problem is a particular instance of maximin peak estimation. Let $p(x) = [p_i(x)]_{i=1}^{N_p}$ be a polynomial vector of objectives. The maximin peak estimation problem is

$$\begin{aligned} P^* &= \max_{t, x_0 \in X_0} \min_i p_i(x) \\ \dot{x}(t) &= f(t, x), \quad t \in [0, T]. \end{aligned} \tag{5.3}$$

Theorem 5.3.1. *The maximin peak estimation problem (5.3) may be upper bounded by a measure program*

$$p^* = \max q \tag{5.4a}$$

$$q + z_i = \langle p_i(x), \mu_p \rangle \quad \forall i = 1..N_p \tag{5.4b}$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \tag{5.4c}$$

$$\mu_0(X_0) = 1 \tag{5.4d}$$

$$q \in \mathbb{R}, z \in \mathbb{R}_+^{N_p} \tag{5.4e}$$

$$\mu, \mu_p \in \mathcal{M}_+([0, T] \times X) \tag{5.4f}$$

$$\mu_0 \in \mathcal{M}_+(X_0). \tag{5.4g}$$

Proof. This is an extension to the measure program (3.2) upper bounding (3.1) with multiple costs. The value q is a lower bound on $\langle p_i, \mu_p \rangle$, and Program (5.4) aims to find the maximum such q . Nonnegative slack variables z_i in (5.4b) fill the gap between the bound q and $\langle p_i, \mu_p \rangle$. \square

Degree- d LMI relaxations provide a decreasing sequence of upper bounds to p^* in (5.4). The Lagrangian of (5.4) is

$$\begin{aligned} L = & \gamma(1 - \langle 1, \mu_0 \rangle) + \sum_{i=1}^{N_p} \alpha_i z_i + \beta_i(q + z_i - \langle p_i, \mu_p \rangle) \\ & + \langle v(t, x), \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu - \mu_p \rangle + q \end{aligned} \quad (5.5)$$

with new dual variables $\beta \in \mathbb{R}^{N_p}$ from constraint (5.4b) and $\alpha \in \mathbb{R}_+^{N_p}$ from the cone constraint $z \in \mathbb{R}_+^{N_p}$. After eliminating α , the dual problem to (5.4) is

$$d^* = \min_{\gamma \in \mathbb{R}} \quad \gamma \quad (5.6a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (5.6b)$$

$$\mathcal{L}_f v(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (5.6c)$$

$$v(t, x) \geq \beta^T p(x) \quad \forall (t, x) \in [0, T] \times X \quad (5.6d)$$

$$v \in C^1([0, T] \times X) \quad (5.6e)$$

$$\beta \in \mathbb{R}_+^{N_p}, \quad \mathbf{1}^T \beta = 1. \quad (5.6f)$$

Strong duality holds between (5.4) and (5.6) by Theorem 3.10 of [17] when $[0, T] \times X$ is compact.

Remark 5.3.1. If a particular term $p_i(x)$ is minimal among $p(x)$ at optimality, then $z_i = 0$ and $\beta_i \neq 0$. The dual variable β is located on an N_p -dimensional simplex, so a single-objective case will feature $\beta = 1$.

5.3.3 Maximin Example

An example of maximin estimation is the following non-autonomous ODE (Example 2.1 from [6]) is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 t - 0.1 x_1 - x_1 x_2 \\ x_1 t - x_2 + x_1^2 \end{bmatrix}. \quad (5.7)$$

Figure 5.3 plots trajectories from equation (5.7) on the initial set $X_0 = \{x \mid (x_1 + 0.75)^2 + x_2^2 = 1\}$ and total set $X = [-3, 2] \times [-2, 2]$. When maximizing $p(x) = x_1$, over the time range $[0, 5]$, the first three bounds are $p_{1:3}^* = [1.5473, 0.4981, 0.4931]$. The second-largest eigenvalue of $\mathbb{M}_1(y) = 2.943 \times 10^{-6}$, so the moment matrix is nearly rank-1 for atom extraction by Algorithm 1. The near-optimal trajectory is displayed in Fig. 5.3a with $x_0^* = [-1.674, -0.383]$ and $x_p^* = [0.493, 0.029]$. With a maximin objective $p(x) = [x_1, x_2]$, the first three bounds are $p_{1:3}^* = [1.0765, 0.3905, 0.3891]$. At $d = 3$, the optimal $\beta = [0.647, 0.353]$ has both elements

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

nonzero, as $p_1(x_p^*) = p_2(x_p^*) = p_3^*$. Fig. 5.4 displays the maximin objective $\min(x_1, x_2)$ along trajectories in Fig. 5.3. x_p^* is reached at time $t_p^* = 2.19$, which is indicated by the blue stars on Fig. 5.4.

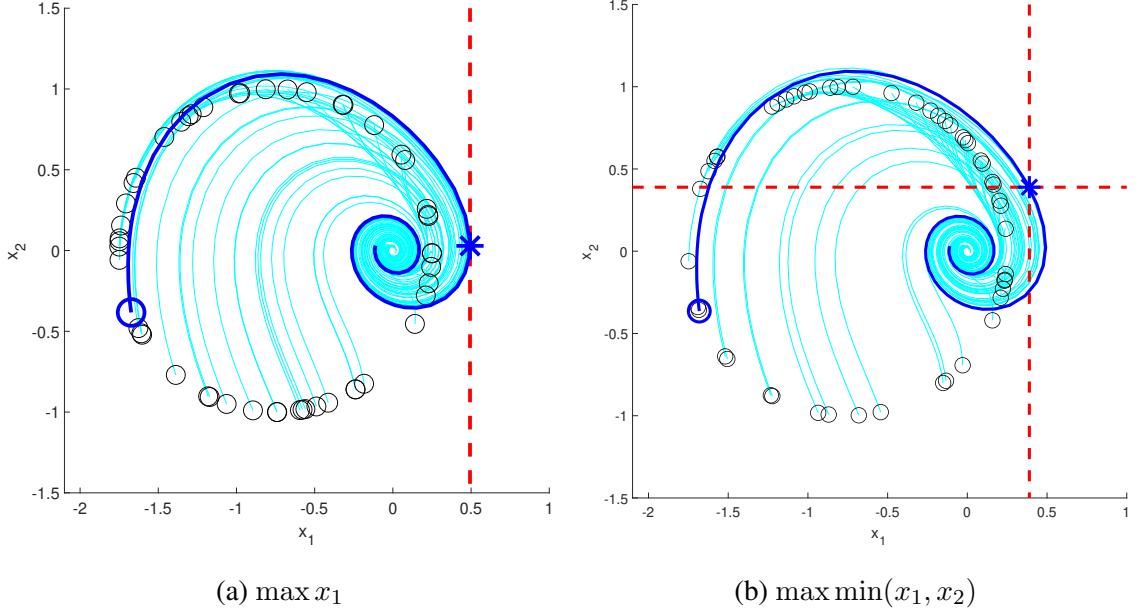


Figure 5.3: Peak analysis of system (5.7) at $d = 3$

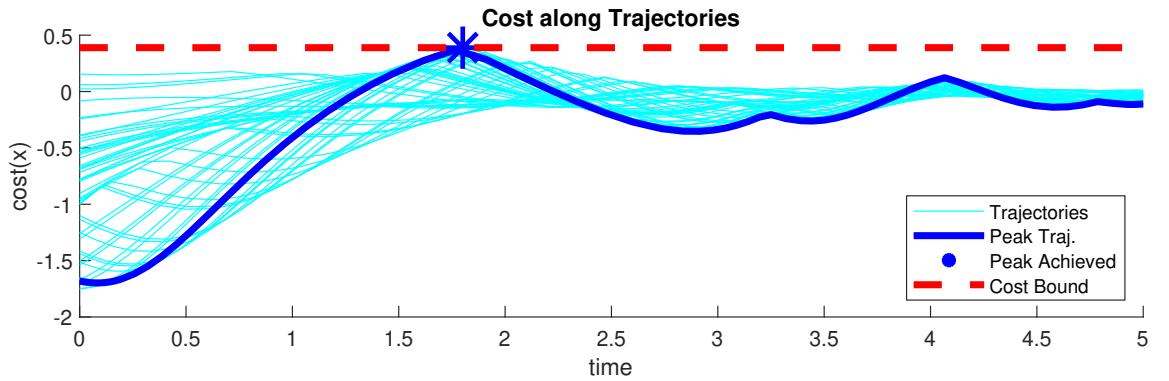


Figure 5.4: The value of $\min(x_1, x_2)$ along trajectories (5.7)

5.3.4 Safety Margin Example and Scaling

The moment-SOS hierarchy can be used to find upper bounds $p_d^* > p^*$ at degrees d . Safety is assured if any upper bound is negative $0 > p_d^* > p^*$. Figure 5.5 visualizes the safety margin for

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

the Flow system (3.9), where the bound of $p^* \leq -0.2831$ was found at the degree-4 relaxation.

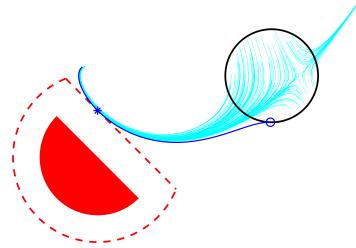


Figure 5.5: Flow system is safe, $p^* \leq -0.2831$

The safety margin of trajectories will generally change if the unsafe set X_u is reparameterized, even in the same coordinate system. Let $q \leq 0$ and $s > 0$ be violation and scaling parameters for the enlarged unsafe set $X_u^{viol} = \{x \mid q \leq 0.5^2 - x_1^2 + (x_2 + 0.7)^2, q \leq -s(x_1 + x_2 - 0.7)\}$. The original unsafe set is $X_u = X_u^{viol}$ with $q = 0$ and $s = \sqrt{2}/2$. Figure 5.6 visualizes contours of regions X_u^q as q decreases from 0 down to -2 for sets with scaling parameters $s = 5$ and $s = 1$. The safety margins of trajectories with respect to X_u will vary as s changes, even as the same set X_u is represented in both cases.

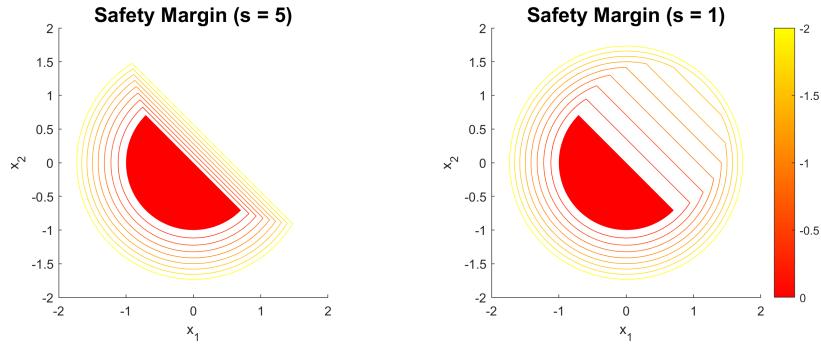


Figure 5.6: Safety margin scaling contours

5.4 Distance Estimation Program

A metric function $c(x, y)$ over the space X with $x, y \in X$ satisfies the following properties [59]:

$$\begin{aligned} c(x, y) &= c(y, x) > 0 & x \neq y \\ c(x, x) &= 0 \\ c(x, y) &\leq c(x, z) + c(z, y) & \forall z \in X \end{aligned}$$

The set of metric functions are closed under addition and pointwise maximums. Every norm $\|\cdot\|$ inspires a metric $c_{\|\cdot\|}(x, y) = \|x - y\|$. The point-set distance function $c(x; Y)$ between a point $x \in X$ and a closed set $Y \subset X$ is defined by:

$$c(x; Y) = \inf_{y \in Y} c(x, y). \quad (5.8)$$

The closest approach as measured by a distance function c that any trajectory takes to the unsafe set X_u in a time horizon of $t \in [0, T]$ can be found by solving

$$\begin{aligned} P^* &= \inf_{t, x_0, y} c(x(t \mid x_0), y) \\ \dot{x}(t) &= f(t, x), \quad t \in [0, T] \\ x(0) &= x_0 \in X_0, \quad y \in X_u. \end{aligned} \quad (5.9)$$

Solving (5.9) requires optimizing over all points $(t, x_0, y) \in [0, T] \times X_0 \times X_u$, which is generically a non-convex and difficult task. Upper bounds to P^* may be found by sampling points (x_0, y) and evaluating $c(x(t \mid x_0), y)$ along these sampled trajectories. Lower bounds to P^* are a universal property of all trajectories, and will satisfy $P^* > 0$ if all trajectories starting from X_0 in the time horizon $[0, T]$ are safe with respect to X_u .

An optimizing trajectory of the Distance program (5.9) may be described by a tuple $\mathcal{T}^* = (y^*, x_0^*, t_p^*)$ using Table 5.1.

Table 5.1: Characterization of optimal trajectory in distance estimation

y^*	location on unsafe set of closest approach
x_0^*	initial condition to produce closest approach
t_p^*	time to reach closest approach from x_0^*

The relationship between these quantities for an optimal trajectory of (5.9) is

$$P^* = c(x(t_p^* \mid x_0^*); X_u) = c(x(t_p^* \mid x_0^*), y^*). \quad (5.10)$$

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Figure 5.7 plots the trajectory of closest approach to X_u in dark blue. This minimal L_2 distance is 0.2831, and the red curve is the level set of all points with a point-set distance 0.2831 to X_u . On the optimal trajectory, the blue circle is $x_0^* \approx (1.489, -0.3998)$, the blue star is $x_p^* = x(t^* | x_0) \approx (0, -0.2997)$, and the blue square is $y^* \approx (-0.2002, -0.4998)$. The closest approach of 0.2831 occurred at time $t^* \approx 0.6180$. Figure 5.8 plots the distance and safety margin contours for the set X_u . These distance contours for a given metric c are independent of the way that X_u is defined (within the same coordinate system).

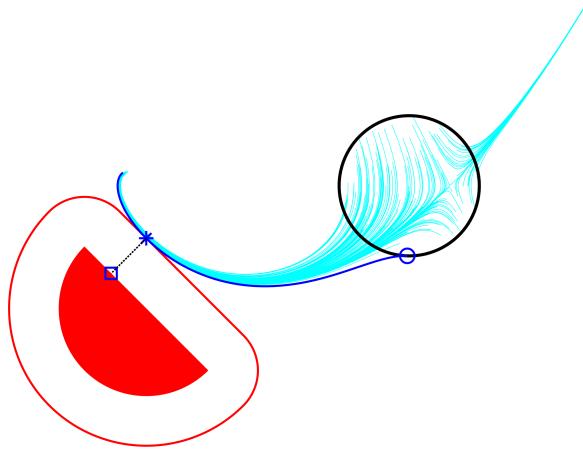


Figure 5.7: Flow system L_2 bound of 0.2831 with optimal trajectory recovery

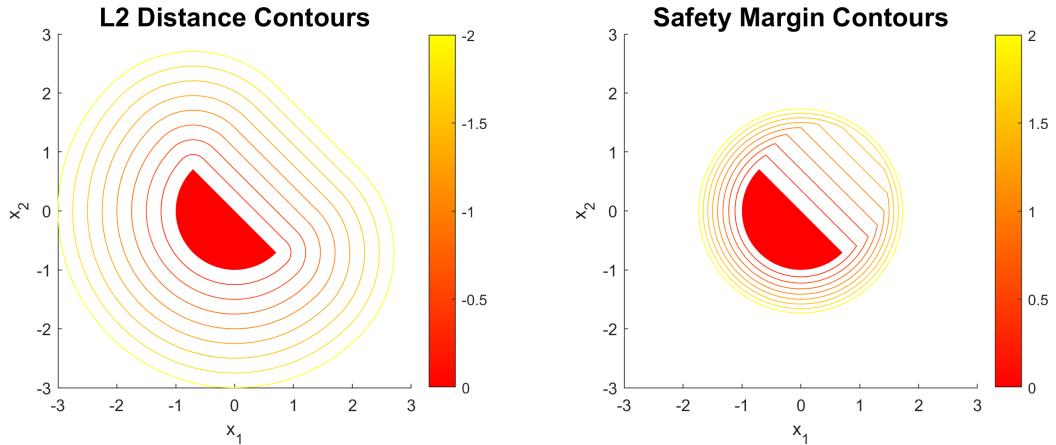


Figure 5.8: Comparison between L_2 distance and safety margin contours

5.4.1 Assumptions

The following assumptions are made in Program (5.9):

- A1 The sets $[0, T]$, X , X_u , X_0 are all compact, $X_0 \subset X$.
- A2 The function $f(t, x)$ is Lipschitz in each argument in the compact set $[0, T] \times X$.
- A3 The cost $c(x, y)$ is C^0 in $X \times X_u$.
- A4 If $x(t \mid x_0) \in \partial X$ for some $t \in [0, T]$, $x_0 \in X_0$, then $x(t' \mid x_0) \notin X \forall t' \in (t, T]$.

A3 relaxes the requirement that c should be a metric, allowing for costs such as $\|x - y\|_2^2$ in addition to the metric $\|x - y\|_2$. The combination of A1 and A3 enforce that $c(x, y)$ is bounded inside $X \times X_u$ by the Weierstrass extreme value theorem. Assumption A4 requires that trajectories leave X immediately after contacting the boundary ∂X .

Remark 5.4.1. A strict ϵ -superset X^ϵ is a set $X^\epsilon \supset X$ in which the boundaries of X^ϵ and X have a positive distance. If trajectories starting in X_0 remain in X at all times $t \in [0, T]$, then any strict ϵ -superset X^ϵ satisfies A4. However, X may not satisfy A4, because there might exist a trajectory remaining in X that is tangent to ∂X .

5.4.2 Measure Program

The problem of $c^* = \min_{(x,y) \in X \times X_u} c(x, y)$ is identical to $c^* = \min_{(x,y) \in X \times X_u} \langle c(x, y), \delta_x \otimes \delta_y \rangle$ for Dirac measures $\delta_x \otimes \delta_y$. The Dirac restriction may be relaxed to minimization over the set of probability measures $c^* = \langle c(x, y), \eta \rangle$, $\eta \in \mathcal{M}_+(X \times X_u)$, $\langle 1, \eta \rangle = 1$ with no change in the objective value c^* . An infinite-dimensional convex LP in measures $(\mu_0, \mu_p, \mu, \eta)$ to bound from below the distance closest approach to X_u starting from X_0 may be developed.

Theorem 5.4.1. Suppose that $f \in C^0$ and A3 holds. Further impose that if $X_0 \subset X$ are both compact then A4 holds. Under these conditions, a lower bound for P^* is,

$$p^* = \inf \langle c(x, y), \eta \rangle \quad (5.11a)$$

$$\pi_\#^x \eta = \pi_\#^x \mu_p \quad (5.11b)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \quad (5.11c)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (5.11d)$$

$$\mu_0 \in \mathcal{M}_+(X_0), \eta \in \mathcal{M}_+(X \times X_u) \quad (5.11e)$$

$$\mu_p, \mu \in \mathcal{M}_+([0, T] \times X) \quad (5.11f)$$

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Proof. Let $\mathcal{T} = (y, x_0, t_p) \in X_u \times X_0 \times [0, T]$ be a tuple representing a trajectory with $x_p = x(t_p | x_0)$ achieving a distance $P = c(x_p, y)$. A set of measures (5.11e)-(5.11f) satisfying constraints (5.11b)-(5.11f) may be constructed from the tuple \mathcal{T} . The initial measure $\mu_0 = \delta_{x=x_0}$, the peak (free-time terminal) measure $\mu_p = \delta_{t=t_p} \otimes \delta_{x=x_p}$ with $x_p = x(t_p | x_0)$, and the joint measure $\eta = \delta_{x_p} \otimes \delta_{y=y}$, are all rank-one atomic probability measures. The measure μ is the occupation measure of $t \mapsto (t, x(t | x_0))$ in times $[0, t_p]$. The distance objective (5.11a) for the tuple \mathcal{T} may be evaluated as,

$$\langle c(x, y), \eta \rangle = \langle c(x, y), \delta_{x=x_p} \otimes \delta_{y=y} \rangle = c(x_p, y) = P. \quad (5.12)$$

The feasible set of (5.11b)-(5.11f) contains all measures constructed from trajectories by the above process, which immediately implies that $p^* \leq P^*$.

□

Remark 5.4.2. As a reminder, the term $\pi_\#^x$ from constraint (5.11b) is the operator performing x -marginalization. Constraint (5.11b) ensures that the x -marginals of η and μ_p are equal: $\forall w \in C(X) : \langle w(x), \eta(x, y) \rangle = \langle w(x), \mu_p(t, x) \rangle$.

We now prove that the measure program in (5.11) has the same objective value as the trajectory program in (5.9) under assumptions A1-A4. In order to accomplish this task, we require a pair of lemmas.

Lemma 5.4.2. Under assumptions A1-A4, the following measure LP has the same optimal value as (5.9):

$$p_c^* = \inf \langle c(x; X_u), \mu_p(t, x) \rangle \quad (5.13a)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \quad (5.13b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (5.13c)$$

$$\mu_0 \in \mathcal{M}_+(X_0), \mu_p, \mu \in \mathcal{M}_+([0, T] \times X) \quad (5.13d)$$

Proof. Problem (5.13) is a peak estimation instance of (3.2) with a continuous (A3) objective of $p(x) = -c(x; X_u)$. Theorem 2.1 of [7] states that the peak estimation LP (5.13) will equal the true peak estimation problem (3.1) (distance estimation problem (5.9)). The measures in (5.13d) contain trajectories that stay within X and terminate on ∂X (agreeing with the non-return assumption A4). □

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Lemma 5.4.3. *Under the assumptions that A1 and A3 hold and that $\nu \in \mathcal{M}_+(X)$ is a probability measure, it follows that*

$$\langle c(x; X_u), \nu(x) \rangle = \inf_{\eta \in \mathcal{M}_+(X \times Y)} \langle c, \eta \rangle : \pi_\#^x \eta = \nu. \quad (5.14)$$

Proof. This follows by Theorem 2.2(a) of [60], given that $X \times X_u$ is compact and c is continuous. \square

Remark 5.4.3. *The parameterized method of [60] assumes that ν has a positive density with respect to the Lebesgue measure on X . However, this assumption of positive density is not required in the statement nor the proof of Theorem 2.2(a) used in [60] (and therefore in Lemma 5.4.3 in this paper).*

Theorem 5.4.4. *Under assumptions A1-A4, $p^* = P^*$.*

Proof. Lemma 5.4.2 states that $p_c^* = P^*$ under assumptions A1-A4. For any solution (μ_p, μ_0, μ) to constraints (5.13b) -(5.13a), Lemma 5.4.3 allows for a measure η to be chosen under $\nu = \pi_\#^x \mu_p$ with cost $\langle c(x; X_u), \pi_\#^x \mu_p(x) \rangle = \langle c, \eta \rangle$. Furthermore, it is not possible to choose an η such that $\langle c(x; X_u), \pi_\#^x \mu_p(x) \rangle \geq \langle c, \eta \rangle$. The infimal objectives $p^* = p_c^*$ are the same, which implies that $p^* = P^*$. \square

5.4.3 Function Program

Dual variables $v(t, x) \in C^1([0, T] \times X)$, $w(x) \in C(X)$, $\gamma \in \mathbb{R}$ over constraints (5.11b)-(5.11d) must be introduced to derive the dual LP to (5.11). The Lagrangian \mathcal{L} of problem (5.11) is:

$$\begin{aligned} \mathcal{L} = & \langle c(x, y), \eta \rangle + \langle v(t, x), \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu - \mu_p \rangle \\ & + \langle w(x), \pi_\#^x \mu_p - \pi_\#^x \eta \rangle + \gamma(1 - \langle 1, \mu_0 \rangle). \end{aligned} \quad (5.15)$$

Recalling that $\forall \eta \in \mathcal{M}_+(X \times Y)$, $w \in C(X)$ the relation that $\langle w(x), \eta(x, y) \rangle = \langle w(x), \pi_\#^x \eta(x) \rangle$ holds, the Lagrangian \mathcal{L} in (5.15) may be reformulated as,

$$\begin{aligned} \mathcal{L} = & \gamma + \langle v(0, x) - \gamma, \mu_0 \rangle + \langle c(x, y) - w(x), \eta \rangle \\ & + \langle w(x) - v(t, x), \mu_p \rangle + \langle \mathcal{L}_f v(t, x), \mu \rangle. \end{aligned} \quad (5.16)$$

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

The dual of program (5.11) is provided by,

$$d^* = \sup_{\gamma, v, w} \inf_{\mu_0, \mu_p, \mu, \eta} \mathcal{L} \quad (5.17a)$$

$$= \sup_{\gamma \in \mathbb{R}} \gamma \quad (5.17b)$$

$$v(0, x) \geq \gamma \quad \forall x \in X_0 \quad (5.17c)$$

$$c(x, y) \geq w(x) \quad \forall (x, y) \in X \times X_u \quad (5.17d)$$

$$w(x) \geq v(t, x) \quad \forall (t, x) \in [0, T] \times X \quad (5.17e)$$

$$\mathcal{L}_f v(t, x) \geq 0 \quad \forall (t, x) \in [0, T] \times X \quad (5.17f)$$

$$w \in C(X) \quad (5.17g)$$

$$v \in C^1([0, T] \times X). \quad (5.17h)$$

Theorem 5.4.5. *Strong duality with $p^* = d^*$ and attainment of optima occurs under assumptions A1-A4.*

Proof. See Appendix A.1. □

Remark 5.4.4. *The continuous function $w(x)$ is a lower bound on the point set distance $c(x; X_u)$ by constraint (5.17d). The auxiliary function $v(t, x)$ is in turn a lower bound on $w(x)$ by constraint (5.17e). This establishes a chain of lower bounds $v(t, x) \leq w(x) \leq c(x; X_u)$ holding $\forall (t, x) \in [0, T] \times X$.*

5.5 Finite-Dimensional Programs

This section presents finite-dimensional SDP truncations to the infinite-dimensional Distance Estimation LPs (5.11) and (5.17).

5.5.1 LMI Approximation

In the case where $c(x, y)$ and $f(t, x)$ are polynomial, (5.11) may be approximated with a converging hierarchy of SDPs. Assume that that X_0 , X , and X_u are Archimedean basic semialgebraic sets, each defined by a finite number of bounded-degree polynomial inequality constraints $X_0 = \{x \mid g_k^0(x) \geq 0\}_{k=1}^{N_0}$, $X = \{x \mid g_k^X(x) \geq 0\}_{k=1}^{N_X}$, and $X_u = \{x \mid g_k^U(x) \geq 0\}_{k=1}^{N_U}$.

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

The polynomial inequality constraints for X_0, X, X_u are of degrees d_k^0, d_k, d_k^U respectively. The Liouville equation in (5.11c) enforces a countably infinite set of linear constraints indexed by all possible $\alpha \in \mathbb{N}^n$, $\beta \in \mathbb{N}$ from (3.6).

Let $(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}, \mathbf{m}^\eta)$ be moment sequences for the measures $(\mu_0, \mu_p, \mu, \eta)$. Define $\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^p)$ as the linear relation induced by (3.6) at the test function $x^\alpha t^\beta$ in terms of moment sequences. The polynomial metric $c(x, y)$ may be expressed as $\sum_{\alpha, \gamma} c_{\alpha\gamma} x^\alpha y^\gamma$ for multi-indices $\alpha, \gamma \in \mathbb{N}^n$. The complexity of dynamics f induces a degree \tilde{d} as $\tilde{d} = d + \lceil \deg(f)/2 \rceil - 1$. The degree- d LMI relaxation of (5.11) with moment sequence variables $(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}, \mathbf{m}^\eta)$ is

$$p_d^* = \min \quad \sum_{\alpha, \gamma} c_{\alpha\gamma} \mathbf{m}_{\alpha\gamma}^\eta. \quad (5.18a)$$

$$\mathbf{m}_{\alpha 0}^\eta = \mathbf{m}_{\alpha 0}^p \quad \forall \alpha \in \mathbb{N}_{\leq 2d}^n \quad (5.18b)$$

$$\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}) = 0 \quad \forall (\alpha, \beta) \in \mathbb{N}_{\leq 2d}^{n+1} \quad (5.18c)$$

$$\mathbf{m}_0^0 = 1 \quad (5.18d)$$

$$\mathbb{M}_d(X_0 \mathbf{m}^0) \succeq 0 \quad (5.18e)$$

$$\mathbb{M}_d(([0, T] \times X) \mathbf{m}^p) \succeq 0 \quad (5.18f)$$

$$\mathbb{M}_{\tilde{d}}(([0, T] \times X) \mathbf{m}) \succeq 0 \quad (5.18g)$$

$$\mathbb{M}_d((X \times X_u) \mathbf{m}^\eta) \succeq 0. \quad (5.18h)$$

Constraints (5.18b)-(5.18d) are finite-dimensional versions of constraints (5.11b)-(5.11d) from the measure LP. In order to ensure convergence $\lim_{d \rightarrow \infty} p_d^* = p^*$, we must establish that all moments of measures are bounded.

Lemma 5.5.1. *The masses of all measures in (5.11) are finite (uniformly bounded) if A1-A4 hold.*

Proof. Constraint (5.11d) imposes that $\langle 1, \mu_0 \rangle = 1$, which further requires that $\langle 1, \mu_p \rangle = \langle 1, \mu_0 \rangle = 1$ by constraint (5.11c) ($v(t, x) = 1$) and $\langle 1, \mu_p \rangle = \langle 1, \eta \rangle = 1$ ($w(x) = 1$). The occupation measure μ likewise has bounded mass with $\langle 1, \mu \rangle = \langle t, \mu^p \rangle < T$ by constraint (5.11c) ($v(t, x) = t$). \square

Lemma 5.5.2. *The measures $(\mu_0, \mu_p, \mu, \eta)$ all have finite moments under Assumptions A1-A4.*

Proof. A sufficient condition for a measure $\tau \in \mathcal{M}_+(X)$ with compact support to be bounded is to have finite mass $\langle 1, \tau \rangle$. In our case, the support of all measures $(\mu_0, \mu_p, \mu, \eta)$ are compact sets by A1. Further, under Assumptions A1-A4, all of these measures have bounded mass (Lemma 5.5.1). This sufficiency is satisfied by all measures $(\mu_0, \mu_p, \mu, \eta)$. \square

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Theorem 5.5.3. *When T is finite and X_0, X, X_u are all Archimedean, the sequence of lower bounds $p_d^* \leq p_{d+1}^* \leq p_{d+2}^* \dots$ will approach p^* as d tends towards ∞ .*

Proof. This convergence is assured by Corollary 8 of [23] under the Archimedean assumption and Lemma 5.5.1. \square

Remark 5.5.1. *Non-polynomial C^0 cost functions $c(x, y)$ may be approximated by polynomials $\tilde{c}(x, y)$ through the Stone-Weierstrass theorem in the compact set $X \times Y$. For every $\epsilon > 0$, there exists a $\tilde{c}(x, y) \in \mathbb{R}[x, y]$ such that $\max_{x \in X, y \in X_u} |c(x, y) - \tilde{c}(x, y)| \leq \epsilon$. Solving the peak estimation problem (5.11) with cost $\tilde{c}(x, y)$ as $\epsilon \rightarrow 0$ will yield convergent bounds to P^* with cost $c(x, y)$. Section 5.9.2 offers an alternative peak estimation problem using polyhedral lifts for costs comprised of the maximum of a set of functions.*

5.5.2 Numerical Considerations

A moment matrix with n variables in degree d has dimension $\binom{n+d}{d}$. The sizes of moment matrices associated with a d relaxation of Problem (5.18) with state $x \in \mathbb{R}^n$, dynamics $f(t, x)$, and induced dynamic degree \tilde{d} , are listed in Table 5.2.

Table 5.2: Sizes of moment matrices in LMI (5.18)

Moment	$\mathbb{M}_d(\mathbf{m}^0)$	$\mathbb{M}_d(\mathbf{m}^p)$	$\mathbb{M}_{\tilde{d}}(\mathbf{m})$	$\mathbb{M}_d(\mathbf{m}^\eta)$
Size	$\binom{n+d}{d}$	$\binom{1+n+d}{d}$	$\binom{1+n+\tilde{d}}{\tilde{d}}$	$\binom{2n+d}{d}$

The computational complexity of solving the SDP formulation of LMI (5.18) scales polynomially as the largest matrix size in Table 5.2, usually $\mathbb{M}_d(\mathbf{m}^\eta)$, except in cases where $f(t, x)$ has a high polynomial degree.

Remark 5.5.2. *The measures μ_p and η may in principle be combined into a larger measure $\tilde{\eta} \in \mathcal{M}_+([0, T] \times X \times X_u)$. The Liouville equation (5.11c) would then read $\pi_\#^{tx}\tilde{\eta} = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu$, and a valid selection of $\tilde{\eta}$ given an optimal trajectory is $\tilde{\eta} = \delta_{t=t_p^*} \otimes \delta_{x=x_p^*} \otimes \delta_{y=y^*}$ with $x_p^* = x(t_p^* \mid x_0^*)$. The measure $\tilde{\eta}$ is defined over $2n + 1$ variables, and the size of its moment matrix at a degree d relaxation is $\binom{1+2n+d}{d}$, as compared to $\binom{2n+d}{d}$ for η . We elected to split up the measures as μ_p and η to reduce the number of variables in the largest measure, and to ensure that the objective (5.11a) is interpretable as an earth-mover distance (from optimal transport literature[53]) between $\pi_\#^x \mu_p$ and a probability distribution over X_u (absorbed into $\pi_\#^x \eta$).*

Remark 5.5.3. *The distance problem (5.9) may also be treated as a peak estimation problem (3.1) with cost $p(x, y) = -c(x, y)$, initial set $X_0 \times X_u$, x -dynamics $\dot{x}(t) = f(t, x(t))$, and y -dynamics $\dot{y}(t) = 0$. The moment matrix $\mathbb{M}_d[\mathbf{m}]$ associated with this peak estimation problem's occupation measure (LMI relaxation of program (3.2)) would have size $\binom{1+2n+\tilde{d}}{d}$. This size is greater than any of the sizes written in Table 5.2.*

Remark 5.5.4. *The atom-extraction-based recovery Algorithm 1 from [32] may be used to approximate near-optimal trajectories if the moment matrices $\mathbb{M}_d(\mathbf{m}^0)$, $\mathbb{M}_d(\mathbf{m}^p)$, and $\mathbb{M}_d(\mathbf{m}^\eta)$ are each low rank. If these matrices are all rank-one, then the near-optimal points (x_p, y, x_0, t_p) may be read directly from the moment sequences $(\mathbf{m}^0, \mathbf{m}^p, \mathbf{m}^\eta)$. The near optimal points from Figure 3.1 were recovered at the degree-4 relaxation of (5.18). The top corner of the moment matrices $\mathbb{M}_d(\mathbf{m}^0)$, $\mathbb{M}_d(\mathbf{m}^p)$, and $\mathbb{M}_d(\mathbf{m}^\eta)$ (containing moments of orders 0-2) have second-largest eigenvalues of 1.87×10^{-5} , 8.82×10^{-6} , 5.87×10^{-7} respectively, as compared to the largest eigenvalues of 3.377, 1.472, 1.380.*

5.5.3 SOS Approximation

The degree- d WSOS truncation of program (5.17) is

$$d_d^* = \max_{\gamma \in \mathbb{R}} \quad \gamma \tag{5.19a}$$

$$v(0, x) - \gamma \in \Sigma[X_0]_{\leq 2d} \tag{5.19b}$$

$$c(x, y) - w(x) \in \Sigma[X \times X_u]_{\leq 2d} \tag{5.19c}$$

$$w(x) - v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2d} \tag{5.19d}$$

$$\mathcal{L}_f v(t, x) \in \Sigma[[0, T] \times X]_{\leq 2\tilde{d}} \tag{5.19e}$$

$$w \in \mathbb{R}[x]_{\leq 2d} \tag{5.19f}$$

$$v \in \mathbb{R}[t, x]_{\leq 2d}. \tag{5.19g}$$

Theorem 5.5.4. *Strong duality holds with $p_k^* = d_k^*$ for all $k \in \mathbb{N}$ between (5.18) and (5.19) under assumptions A1-A5.*

Proof. Refer to Corollary 8 of [23] (Archimedean condition and bounded masses), as well as to the proof of Theorem 4 and Lemma 4 in Appendix D of [21]. \square

5.6 Exploiting Correlative Sparsity

Many costs $c(x, y)$ exhibit an additively separable structure such that c can be decomposed into the sum of new terms $c(x, y) = \sum_i c_i(x_i, y_i)$. Each term c_i in the sum is a function purely of (x_i, y_i) . Examples include the L_p family of distance functions, such as the squared L_2 cost $c(x, y) = \sum_i (x_i - y_i)^2$. The theory of Correlative Sparsity in polynomial optimization, briefly reviewed below, can be used to substantially reduce the computational complexity entailed in solving the distance estimation SDPs when c is additively separable [61]. This decomposition does not require prior structure on the set $X \times X_u$. Other types of reducible structure (if applicable) include Term Sparsity [62], symmetry [63], and network dynamics [64]. These forms of structure may be combined if present, such as in Correlative and Term Sparsity [65].

5.6.1 Correlative Sparsity Background

Let $\mathbb{K} = \{x \mid g_k(x) \geq 0, k = 1, \dots, N\}$ be an Archimedean basic semialgebraic set and $\phi(x)$ be a polynomial. The Correlative Sparsity Pattern (CSP) associated to $(\phi(x), g)$ is a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with vertices \mathcal{V} and edges \mathcal{E} . Each of the n vertices in \mathcal{V} corresponds to a variable x_1, \dots, x_n . An edge $(x_i, x_j) \in \mathcal{E}$ appears if variables x_i and x_j are multiplied together in a monomial in $\phi(x)$, or if they appear together in at least one constraint $g_k(x)$ [61].

The correlative sparsity pattern of $(\phi(x), g)$ may be characterized by sets I of variables and sets J of constraints. The p sets I should satisfy the following two properties:

1. (Coverage) $\bigcup_{j=1}^p I_j = \mathcal{V}$
2. (Running Intersection Property) For all $k = 1, \dots, p-1$: $I_{k+1} \cap \bigcup_{j=1}^k I_j \subseteq I_s$ for some $s \leq k$

Equivalently, the sets I are the maximal cliques of a chordal extension of $\mathcal{G}(\mathcal{V}, \mathcal{E})$ [66]. The sets $J = \{J_i\}_{i=1}^p$ are a partition over constraints $g_k(x) \geq 0$. The number k is in J_i for $k = 1, \dots, N_X$ if all variables involved in the constraint polynomial $g_k(x)$ are contained within the set I_i . Let the notation $x(I_i)$ denote the variables in x that are members of the set I_i . A sufficient sparse representation of positivity certificates may be developed for $(\phi(x), g)$ satisfying an admissible correlative sparsity pattern (I, J) [67]:

$$\begin{aligned} \phi(x) &= \sum_{i=1}^p \sigma_{i0}(x(I_i)) + \sum_{k \in J_i} \sigma_k(x(I_i)) g_k(x) \\ \sigma_{i0}(x) &\in \Sigma[x(I_i)] \quad \sigma_k(x) \in \Sigma[x(I_i)] \quad \forall i = 1, \dots, p. \end{aligned} \tag{5.20}$$

Equation (5.20) is a sparse version of the Putinar certificate in (2.13). The sparse certificate (5.20) is additionally necessary for the \mathcal{G} -sparse polynomial $\phi(x)$ to be positive over \mathbb{K} if (I, J) satisfies the Running Intersection Property and a sparse Archimedean property holds: that there exist finite constants $R_i > 0$ for $i = 1..n$ such that $R_i^2 - \|x(I_i)\|_2^2$ is in the quadratic module (2.11) of constraints $Q[\{g_k\}_{k \in J_i}]$ [67].

5.6.2 Correlative Sparsity for Distance Estimation

Constraint (5.17d) will exhibit correlative sparsity when $c(x, y)$ is additively separable:

$$\sum_{i=1}^n c_i(x_i, y_i) - w(x) \geq 0 \quad \forall (x, y) \in X \times X_u. \quad (5.21)$$

The product-structure support set of Equation (5.21) may be expressed as

$$X \times X_u = \{(x, y) \mid g_1(x) \geq 0, \dots, g_{N_X}(x) \geq 0, \\ g_{N_X+1}(y) \geq 0, \dots, g_{N_X+N_U}(y) \geq 0\}. \quad (5.22)$$

The correlative sparsity graph of (5.21) is the graph Cartesian product of the complete graph K_n by K_2 and is visualized at $n = 4$ by the nodes and black lines in Figure 5.9. Black lines imply that there is a link between variables. The black lines are drawn between each pair (x_i, y_i) from the cost terms c_i . The polynomial $w(x)$ involves mixed monomials of all variables $(x) = (x_1, x_2, x_3, x_4)$. Prior knowledge on the constraints of X_u are not assumed in advance, so the variables are $(y) = (y_1, y_2, y_3, y_4)$ joined together. A CSP (I, J) associated with this system is

$$\begin{array}{ll} I_1 = \{x_1, x_2, x_3, x_4, y_1\} & J_1 = \{1, \dots, N_X\} \\ I_2 = \{x_2, x_3, x_4, y_1, y_2\} & J_2 = \emptyset \\ I_3 = \{x_3, x_4, y_1, y_2, y_3\} & J_3 = \emptyset \\ I_4 = \{x_4, y_1, y_2, y_3, y_4\} & J_4 = \{N_X + 1, \dots, N_X + N_U\}. \end{array}$$

Figure 5.9 illustrates a chordal extension of the CSP graph with new edges displayed as red dashed lines. These new edges appear by connecting all variables in I_1 together in a clique and by following a similar process for I_2, \dots, I_4 .

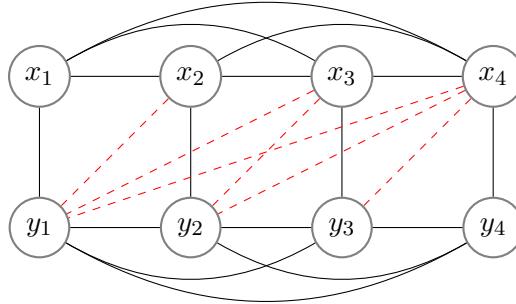


Figure 5.9: CSP with 4-States and Chordal Extension

For a unsafe distance bounding problem with a additively separable $c(x, y) = \sum_i c(x_i, y_i)$ with n states, the CSP (I, J) is

$$\begin{aligned} I_1 &= \{x_1, \dots, x_n, y_1\} & J_1 &= \{1, \dots, N_X\} \\ I_i &= \{x_i, \dots, x_n, y_1, \dots, y_i\} & J_i &= \emptyset, \quad \forall i = 2, \dots, n-1 \\ I_n &= \{x_n, y_1, \dots, y_n\} & J_n &= \{N_X + 1, \dots, N_X + N_U\}. \end{aligned} \quad (5.23)$$

A total of $(n-1)n/2$ new edges are added in the chordal extension. Letting $y_{1:i}$ be the collection of variables (y_1, y_2, \dots, y_i) for an index $i \in 1..n$ (and with a similar definition for $x_{i:n}$), a correlative sparse certificate of positivity for constraint (5.17d) is

$$\begin{aligned} \sum_{i=1}^n c_i(x_i, y_i) - w(x) &= \sum_{i=1}^n \sigma_{i0}(x_{i:n}, y_{1:i}) + \sum_{k=1}^{N_X} \sigma_k(x, y_1) g_k(x) \\ &\quad + \sum_{k=N_X+1}^{N_X+N_U} \sigma_k(x_n, y) g_k(y). \end{aligned} \quad (5.24)$$

with sum-of-squares multipliers

$$\begin{aligned} \sigma_{i0}(x, y) &\in \Sigma[x_{i:n}, y_{1:i}] & \forall i = 1, \dots, p \\ \sigma_k(x, y) &\in \Sigma[x, y_1] & \forall k = 1, \dots, N_X \\ \sigma_k(x, y) &\in \Sigma[x_n, y] & \forall k = N_X + 1, \dots, N_X + N_U. \end{aligned} \quad (5.25)$$

The application of correlative sparsity to the distance problem replaces constraint (5.19c) by (5.24).

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Remark 5.6.1. *The CSP decomposition in (5.23) is nonunique. As an example, the following decompositions are all valid for $n = 3$ (satisfying the Running Intersection Property):*

$$\begin{array}{ll} I_1 = \{x_1, x_2, x_3, y_1\} & I'_1 = \{x_1, x_2, x_3, y_3\} \\ I_2 = \{x_2, x_3, y_1, y_3\} & I'_2 = \{x_1, x_2, y_2, y_3\} \\ I_3 = \{x_2, y_1, y_2, y_3\} & I'_3 = \{x_1, y_1, y_2, y_3\} \end{array}$$

The original constraint (5.17d) is dual to the joint measure $\eta \in \mathcal{M}_+(X \times Y)$. Correlative sparsity may be applied to the measure program by splitting η into new measures $\eta_1 \in \mathcal{M}_+(\mathbb{R} \times \mathbb{R})$, $\eta_n \in \mathcal{M}_+(\mathbb{R} \times X_u)$ and $\eta_i \in \mathcal{M}_+(\mathbb{R}^{n+1})$ for $i = 2, \dots, n-1$, following the procedure in [67]. These measures will align on overlaps with $\pi_{\#}^{I_i \cap I_{i+1}} \eta_i = \pi_{\#}^{I_i \cap I_{i+1}} \eta_{i+1}$, $\forall i = 1, \dots, n-1$. At a degree d relaxation, the moment matrix of η in (5.18) has size $\binom{2n+d}{d}$. Each of the n moment matrices of $\{\eta_i\}_{i=1}^n$ has a size of $\binom{n+1+d}{d}$. For example, a problem with $n = 4, d = 4$ will have a moment matrix for η of size $\binom{12}{4} = 495$, while the moment matrices for each of the $\eta_{(1:4)}$ are of size $\binom{9}{4} = 126$.

5.7 Shape Safety

The distance estimation problem may be extended to sets or shapes travelling along trajectories, bounding the minimum distance between points on the shape and the unsafe set. An example application is in quantifying safety of rigid body dynamics, such as finding the closest distance between all points on an airplane and points on a mountain.

5.7.1 Shape Safety Background

Let $X \subset \mathbb{R}^n$ be a region of space with unsafe set X_u , and $c(x, y)$ be a distance function. The state $\omega \in \Omega$ (such as position and angular orientation) follows dynamics $\dot{\omega}(t) = f(t, \omega)$ between times $t \in [0, T]$. A trajectory is $\omega(t | \omega_0)$ for some initial state $\omega_0 \in \Omega_0 \subset \Omega$. The shape of the object is a set S . There exists a mapping $A(s; \omega) : S \times \Omega \rightarrow X$ that provides the transformation between local coordinates on the shape (s) to global coordinates in X .

Examples of a shape traveling along trajectories are detailed in Figure 5.10. The shape $S = [-0.1, 0.1]^2$ is the pink square. The left hand plot is a pure translation after a $5\pi/12$ radian rotation, and the right plot involves a rigid body transformation.

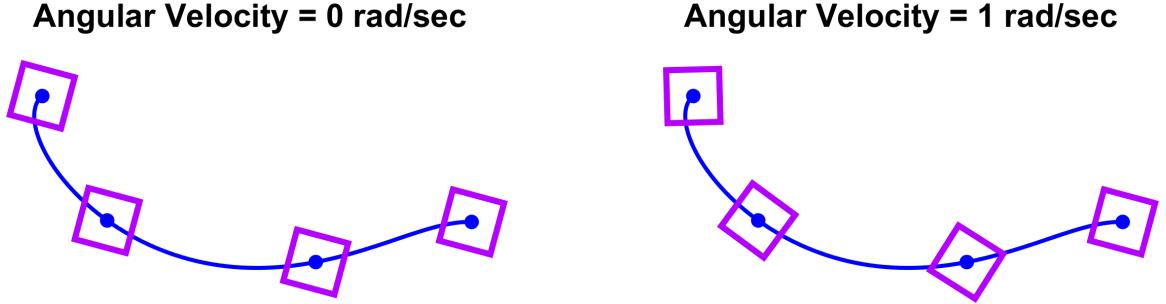


Figure 5.10: Shape moving and rotating along Flow (3.9) trajectories

The distance estimation task with shapes is to bound,

$$\begin{aligned} P^* &= \inf_{t, \omega_0 \in \Omega_0, s \in S, y \in X_u} c(A(s; \omega(t | \omega_0)), y) \\ \dot{\omega}(t) &= f(t, \omega), \quad \forall t \in [0, T]. \end{aligned} \tag{5.26}$$

For each trajectory in state $\omega(t | \omega_0)$, problem (5.26) ranges over all points in the shape $s \in S$ and points in the unsafe set $y \in X_u$ to find the closest approach. An optimal trajectory of the shape distance program may be expressed as $\mathcal{T}_s^* = (y^*, s^*, \omega_0^*, t_p^*)$ with $\omega_p^* = \omega(t_p^* | \omega_0^*)$, $x_p^* = A(s^*; \omega_p^*)$ and

$$P^* = c(A(s^*; \omega_p^*), X_u) = c(A(s^*; \omega(t_p^* | \omega_0^*)), y^*).$$

Remark 5.7.1. The objective in (5.26) can be expressed using

$$c_A(\omega; S, X_u) = \inf_{s \in S, y \in X_u} c(A(s; \omega), y) \tag{5.27}$$

as $c_A(\omega(t | \omega_0); S, X_u)$.

5.7.2 Assumptions

The following assumptions are made in the Shape Distance program (5.26):

A1' The sets $[0, T]$, Ω , S , X , X_u are compact and $\Omega_0 \subset \Omega$.

A2' The function $f(t, \omega)$ is Lipschitz in each argument.

A3' The cost $c(x, y)$ is C^0 .

A4' The coordinate transformation function $A(s; \omega)$ is C^0 .

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

A5' If $\omega(t \mid \omega_0) \in \partial\Omega$ for some $t \in [0, T]$, $\omega_0 \in \Omega_0$, then $\omega(t \mid \omega_0) \notin \Omega \forall t' \in (t, T]$.

A6' If $\exists s \in S$ such that $A(s; \omega(t \mid \omega_0)) \notin X$ or $A(s; \omega(t \mid \omega_0)) \in \partial X$ for some $t \in [0, T]$, $\omega_0 \in \Omega_0$, then $A(s; \omega(t' \mid \omega_0)) \notin X \forall t' \in (t, T]$.

An alternative assumption used instead of A5'-A6' is that $\omega(t \mid \Omega_0)$ stays in Ω for all $\omega_0 \in \Omega_0$ and $A(s; \omega(t \mid \omega_0)) \in X$ for all $s \in S, t \in [0, T]$.

5.7.3 Shape Distance Measure Program

Program (5.26) involves a distance objective $c(x, y)$, where the point $x = A(s; \omega)$ is given by a coordinate transformation between body coordinates s and the evolving orientation ω . In order to formulate a measure program to (5.26), a shape measure $\mu_s \in \mathcal{M}_+(S \times \Omega)$ may be added to bridge the gap between the changing orientation $\dot{\omega}$ and the comparison distance x . The shape measure contains information on the orientation ω and body coordinate s that yields the closest point x ,

$$\langle z(\omega), \mu_p(t, \omega) \rangle = \langle z(\omega), \mu_s(s, \omega) \rangle \quad \forall z \in C(\Omega) \quad (5.28a)$$

$$\langle w(x), \eta(x, y) \rangle = \langle w(A(s; \omega)), \mu_s(s, \omega) \rangle \quad \forall w \in C(X). \quad (5.28b)$$

The shape measure μ_s chooses the worst-case body coordinate s and orientation ω from μ_p (5.28a), such that the point $x = A(s; \omega)$ comes as close as possible to the unsafe set's coordinate y (5.28b). We retain the coordinate x in order to decrease the computational complexity of the SDPs, as elaborated upon further in Remark 5.5.2.

The infinite dimensional measure program that lower bounds (5.26) is,

$$p^* = \inf \langle c(x, y), \eta \rangle \quad (5.29a)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \quad (5.29b)$$

$$\pi_\#^\omega \mu_p = \pi_\#^\omega \mu_s \quad (5.29c)$$

$$\pi_\#^x \eta = A(s; \omega)_\# \mu_s \quad (5.29d)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (5.29e)$$

$$\mu_0 \in \mathcal{M}_+(\Omega_0), \eta \in \mathcal{M}_+(X \times X_u) \quad (5.29f)$$

$$\mu_s \in \mathcal{M}_+(\Omega \times S) \quad (5.29g)$$

$$\mu_p, \mu \in \mathcal{M}_+([0, T] \times \Omega). \quad (5.29h)$$

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Constraint (5.11b) in the original distance formulation is now split between (5.29c) and (5.29d) (which are equivalent to (5.28b) and (5.28a)). Problem (5.29) inherits all convergence and duality properties of the original (5.11) under the appropriately modified set of assumptions A1'-A6'.

Theorem 5.7.1. *Under A3'-A4' (and additionally A5'-A6' when all sets in A1' are compact possibly excluding $[0, T]$), the Shape programs (5.26) and (5.29) are related by $p^* \leq P^*$.*

Proof. This proof will follow the same pattern as Theorem 5.4.1's proof. A set of measures that are feasible solutions for the constraints of (5.29) may be constructed for any trajectory $\mathcal{T}_s = (y, s, \omega_0, t_p)$ with $\omega_p = \omega(t_p | \omega_0)$, $x_p = A(s; \omega_p)$. One choice of these measures are $\mu_0 = \delta_{\omega=\omega_0}$, $\mu_p = \delta_{t=t_p} \otimes \delta_{\omega=\omega_p}$, $\eta = \delta_{x=x_p} \otimes \delta_{y=y}$, $\mu_s = \delta_{s=s} \otimes \delta_{\omega=\omega_p}$ and μ as the occupation measure $t \mapsto (t, \omega(t | \omega_0^*))$ in times $[0, t_p^*]$. The feasible set of the constraints contains all trajectory-constructed measures, so $p^* \leq P^*$. \square

Lemma 5.7.2. *All measures in (5.29) have bounded mass under Assumption A1'.*

Proof. This follows from the steps of Lemma 5.5.1. The conditions hold that $1 = \langle 1, \mu_0 \rangle = \langle 1, \mu_p \rangle$ (5.29b), $\langle 1, \mu_p \rangle = \langle 1, \mu_s \rangle$ (5.29c), $\langle 1, \mu_s \rangle = \langle 1, \eta \rangle$ (5.29d), and $\langle 1, \mu \rangle \leq T$ by (5.29b). \square

Lemma 5.7.3. *The following peak estimation problem has the same optimal value as (5.26) under A1'-A6':*

$$p_c^* = \inf \langle c_A(\omega; S, X_u), \mu_p(t, \omega) \rangle \quad (5.30a)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu \quad (5.30b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (5.30c)$$

$$\mu_0 \in \mathcal{M}_+(\Omega_0), \eta \in \mathcal{M}_+(X \times X_u) \quad (5.30d)$$

$$\mu_s \in \mathcal{M}_+(\Omega \times S) \quad (5.30e)$$

$$\mu_p, \mu \in \mathcal{M}_+([0, T] \times \Omega). \quad (5.30f)$$

Proof. Refer to the proof of Lemma 5.4.2, with a shape-objective from (5.27). \square

Theorem 5.7.4. *Under A1'-A6', the optimal values of (5.29) and (5.26) are equal ($P^* = p^*$).*

Proof. This proof repeats same process used in Theorem 5.4.4. Lemma 5.7.3 is used in place of Lemma 5.4.2. The reasoning of Lemma 5.4.3 is employed to construct infima-agreeing measures μ_s, η given a μ_p from (5.30f) consistent with the marginal constraints (5.29c) and (5.29d). \square

5.7.4 Shape Distance Function Program

Defining a new dual function $z(\omega)$ against constraint (5.29c) (also observed in (5.28a)), the Lagrangian of problem (5.29) is:

$$\begin{aligned}\mathcal{L} = & \langle c(x, y), \eta \rangle + \langle v(t, x), \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu - \mu_p \rangle \\ & + \langle z(\omega), \pi_\#^\omega (\mu_p - \mu_s) \rangle + \gamma(1 - \langle 1, u_0 \rangle) \\ & + \langle w(x), A(s; \omega)_\# \mu_s - \pi_\#^x \eta \rangle.\end{aligned}\quad (5.31)$$

The Lagrangian in (5.31) may be manipulated into,

$$\begin{aligned}\mathcal{L} = & \gamma + \langle c(x, y) - w(x), \eta \rangle + \langle v(0, \omega) - \gamma, \mu_0 \rangle \\ & + \langle \mathcal{L}_f v(t, \omega), \mu \rangle + \langle z(\omega) - v(t, \omega), \mu_p \rangle \\ & + \langle w(A(s; \omega)) - z(\omega), \mu_s \rangle.\end{aligned}\quad (5.32)$$

The dual of program (5.29) provided by minimizing the Lagrangian (5.32) with respect to $(\eta, \mu_s, \mu_p, \mu, \mu_0)$ is,

$$d^* = \sup_{\gamma \in \mathbb{R}} \quad \gamma \quad (5.33a)$$

$$v(0, \omega) \geq \gamma \quad \forall x \in \Omega_0 \quad (5.33b)$$

$$c(x, y) \geq w(x) \quad \forall (x, y) \in X \times X_u \quad (5.33c)$$

$$w(A(s; \omega)) \geq z(\omega) \quad \forall (s, \omega) \in S \times \Omega \quad (5.33d)$$

$$z(\omega) \geq v(t, \omega) \quad \forall (t, \omega) \in [0, T] \times \Omega \quad (5.33e)$$

$$\mathcal{L}_f v(t, \omega) \geq 0 \quad \forall (t, \omega) \in [0, T] \times \Omega \quad (5.33f)$$

$$w \in C(X), z \in C(\Omega) \quad (5.33g)$$

$$v \in C^1([0, T] \times X). \quad (5.33h)$$

Theorem 5.7.5. *Problems (5.29) and (5.33) are strongly dual under assumptions A1'-A6'.*

Proof. This holds by extending the proof of Theorem 5.4.5 found in Appendix A.1 and applying Theorem 2.6 of [40]. \square

Remark 5.7.2. *Program 5.33 imposes that a chain of lower bounds $v(t, \omega) \leq z(\omega) \leq w(A(s; \omega)) \leq c(A(s; \omega), y)$ holds for all $(s, \omega, t, y) \in S \times \Omega \times [0, T] \times X_u$ (similar in principle to Remark 5.4.4).*

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

Remark 5.7.3. We briefly note that the LMI formulation of (5.29) will converge to P^* under assumptions A1'-A6' if all sets $[0, T]$, X , X_u , Ω_0 , Ω , S are Archimedean and if $f(t, \omega) \in \mathbb{R}[t, \omega]$, $A(s; \omega) \in \mathbb{R}[s, \omega]$ (from Theorem 5.5.3). Constraint (5.28b) induces a linear expression in moments for (η, μ_s) for each $\alpha \in \mathbb{N}^n$: $\langle x^\alpha, \eta \rangle = \langle A(s; \omega)^\alpha, \mu_s \rangle$.

Remark 5.7.4. If $A(s; \omega)$ is polynomial with degree κ , then the d -degree relaxation of problem (5.29) involves moments of μ_s up to order $2\kappa d$. For a system with N_ω orientation states and N_s shape variables, the size of the moment matrix for μ_s is then $\binom{N_s + N_\omega + \kappa d}{\kappa d}$. LMI constraints associated with μ_s can become bottlenecks to computation, surpassing the contributions of μ and η as k increases.

Remark 5.7.5. Continuing the discussion Remark 5.5.2, the measures $\mu_s(s, \omega)$ and $\eta(x, y)$ may be combined together into a larger measure $\eta_s(s, \omega, y) \in \mathcal{M}_+(S \times \Omega \times X_u)$ with objective $\inf \langle c(A(s; \omega), y), \eta_s \rangle$ and constraint $\pi_\#^\omega \mu_p = \pi_\#^\omega \eta_s$. The moment matrix for η_s would have the generally intractable size $\binom{N_s + N_\omega + n + \kappa d}{\kappa d}$.

5.8 Numerical Examples

All code was written in MATLAB 2021a, and is publicly available at the link <https://github.com/Jarmill/distance>. The SDPs were formulated by Gloptipoly3 [30] through a Yalmip interface [68], and were solved using Mosek [49]. The experimental platform was an Intel i9 CPU with a clock frequency of 2.30 GHz and 64.0 GB of RAM. The squared- L_2 cost $c(x, y) = \sum_i (x_i - y_i)^2$ is used in solving Problem (5.18) unless otherwise specified. The documented bounds are the square roots of the returned quantities, yielding lower bounds to the L_2 distance.

5.8.1 Flow System with Moon

The half-circle unsafe set in Figure 5.8 is a convex set. The moon-shaped unsafe set X_u in Figure 5.11 is the nonconvex region outside the circle with radius 1.16 centered at $(0.6596, 0.3989)$ and inside the circle with radius 0.8 centered at $(0.4, -0.4)$. The dotted red line demonstrates that trajectories of the Flow system would be deemed unsafe if X_u was relaxed to its convex hull.

The L_2 distance bound of 0.1592 in Figure 5.12 was found at the degree-5 relaxation of Problem (5.18) with $X = [-3, 3]^2$. The moment matrices $\mathbb{M}_d(m^0)$, $\mathbb{M}_d(m^p)$, $\mathbb{M}_d(m^\eta)$ at $d = 5$ were approximately rank-1 and near-optimal trajectories were successfully extracted. This near-optimal trajectory starts at $x_0^* \approx (1.489, -0.3998)$ and reaches a closest distance between $x_p^* \approx$

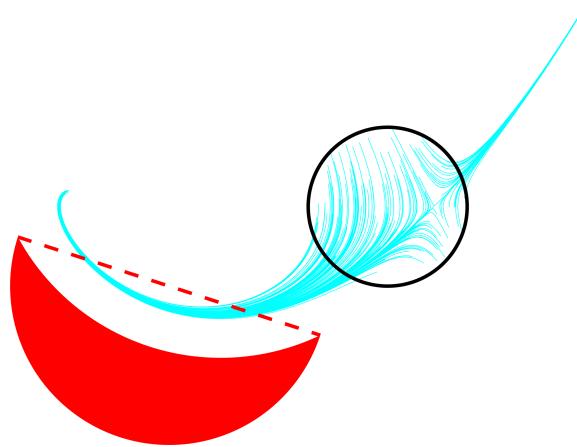


Figure 5.11: Collision if X_u is relaxed to its convex hull.

$(1.113, -0.4956)$ and $y^* \approx (1.161, -0.6472)$ at time $t_p^* \approx 0.1727$. The distance bounds computed at the first five relaxations are $L_2^{1:5} = [1.487 \times 10^{-4}, 2.433 \times 10^{-4}, 0.1501, 0.1592, 0.1592]$.

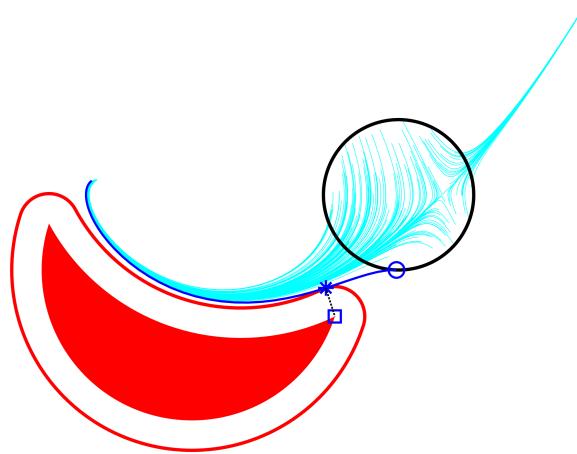


Figure 5.12: Moon unsafe-set has an L_2 bound of 0.1592

5.8.2 Twist System

This subsection performs peak estimation on the Twist system from [69]:

$$\dot{x}_i(t) = \sum_j B_{ij}^1 x_j(t) - B_{ij}^3 (4x_j^3(t) - 3x_j(t))/2, \quad (5.34)$$

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

under the choices of parameters

$$B^1 = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix} \quad B^3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \quad (5.35)$$

Choosing different parameter matrices B^1 (linear) and B^3 (cubic) yields a family of dynamical systems, many of which are attractors, and some of which possess limit cycles. We note that the Twist system possesses a symmetry $x \leftrightarrow -x$.

The cyan curves in each panel of Figure 5.13 are plots of trajectories of the Twist system between times $t \in [0, 5]$. These trajectories start at the $X_0 = \{x \mid (x_1 + 0.5)^2 + x_2^2 + x_3^2 \leq 0.2^2\}$ which is pictured by the grey spheres. The unsafe set $X_u = \{x \mid (x_1 - 0.25)^2 + x_2^2 + x_3^2 \leq 0.2^2, x_3 \leq 0\}$ is drawn in the red half-spheres. The underlying space is $X = [-1, 1]^3$.

The red shell in Figure 5.13a is the cloud of points within an L_2 distance of 0.0427 of X_u , as found through a degree 5 relaxation of (5.18). Figure 5.13b involves an L_4 contour of 0.0411, also found at order 5. The first few distance bounds for the L_2 distance are $L_2^{1:5} = [0, 0, 0.0336, 0.0425, 0.0427]$, and for the L_4 distance are $L_4^{2:5} = [0, 0.0298, 0.0408, 0.0413]$. Fourth degree moments are required for the L_4 metric, so the $L_4^{2:5}$ sequence starts at order 2.

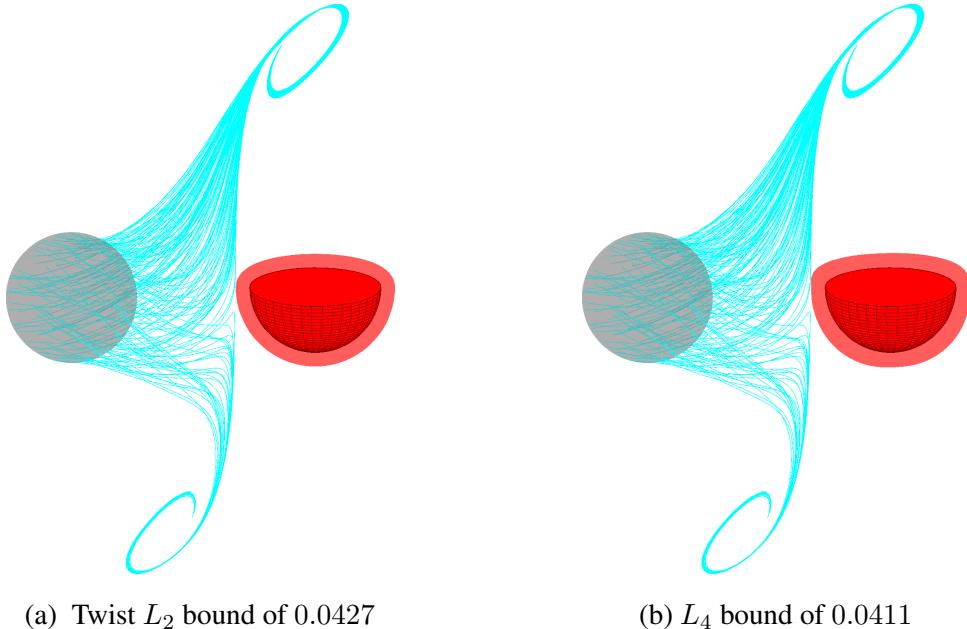


Figure 5.13: Distance contours at order-5 relaxation for the Twist system (5.34)

Table 5.3 and 5.4 lists the L_2 bounds and runtimes respectively generated by a distance estimation task between trajectories and the half sphere of the above L_2 Twist system example. The high-degree relaxations (orders 4 and 5) are significantly faster as found by solving the SDP associated with the sparse LMI (dual to the sparse SOS with Putinar expression (5.24)) as compared to the standard program (5.18). The certifiable L_2 bounds returned are roughly equivalent between relaxations.

Table 5.3: L_2 bounds for the Twist Example

order	2	3	4	5	6
Standard LMI (5.18)	0.000	0.0313	0.0425	0.0429	0.0429
Sparse LMI with (5.24)	0.000	0.0311	0.0424	0.0430	0.0429

Table 5.4: Time in seconds for the Twist Example

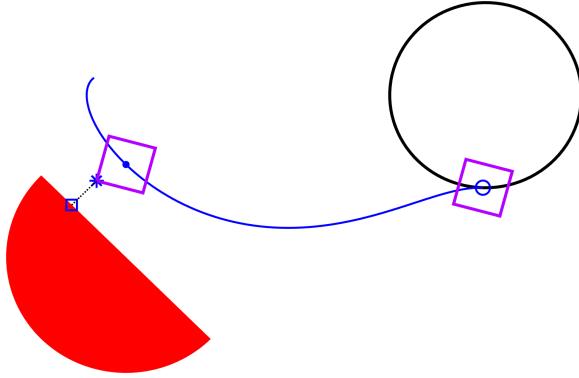
order	2	3	4	5	6
Standard LMI (5.18)	0.32	1.92	47.55	502.29	4028.94
Sparse LMI with (5.24)	0.31	1.19	7.07	45.89	184.42

5.8.3 Shape Examples

Figure 5.14 visualizes a near-optimal trajectory of the shape distance estimation for orientations $\varphi \in \mathbb{R}^2$ evolving as the flow system with an initial condition $\Omega_0 = \{\varphi : (\varphi_1 - 1.5)^2 + \varphi_2^2 \leq 0.4^2\}$ in the space $\Omega : (\varphi_1, \varphi_2) \in [-3, 3]^2$, $\varphi_3^2 + \varphi_4^2 = 1$ (with a state set of $X = [-3, 3]^2$). Suboptimal trajectories were suppressed in visualization to highlight the shape structure and attributes of the near-optimal trajectory. The degree-1 coordinate transformation function A for pure translation with a constant rotation of $5\pi/12$ is

$$A(s; \varphi) = \begin{bmatrix} \cos(5\pi/12)s_1 - \sin(5\pi/12)s_2 + \varphi_1 \\ \cos(5\pi/12)s_1 + \sin(5\pi/12)s_2 + \varphi_2 \end{bmatrix}. \quad (5.36)$$

This near-optimal trajectory with an L_2 distance bound of 0.1465 was found at a degree-4 relaxation of Problem (5.29). The near-optimal trajectory is described by $\varphi_0^* \approx (1.489, -0.3887)$, $t_p^* \approx 3.090$, $\varphi_p^* \approx (-0.1225, -0.3704)$, $s^* \approx (-0.1, 0.1)$, $x_p^* \approx (0, -0.2997)$, and $y^* \approx (-0.2261, -0.4739)$. The first five distance bounds are $L_2^{1:5} = [1.205 \times 10^{-4}, 4.245 \times 10^{-4}, 0.1424, 0.1465, 0.1465]$.


 Figure 5.14: Translation, L_2 bound of 0.1465

In the following example, the shape S is now rotating at an angular velocity of 1 radian/second, as shown in the right panel of Fig. 5.10. The orientation $\varphi \in SE(2)$ may be expressed as a semialgebraic lift through $\varphi \in \mathbb{R}^4$ with trigonometric terms $\varphi_3^2 + \varphi_4^2 = 1$. The dynamics for this system are

$$\dot{\varphi} = \begin{bmatrix} \varphi_2 & -\varphi_1 - \varphi_2 + \frac{1}{3}\varphi_1^3 & -\varphi_4 & \varphi_3 \end{bmatrix}^T. \quad (5.37)$$

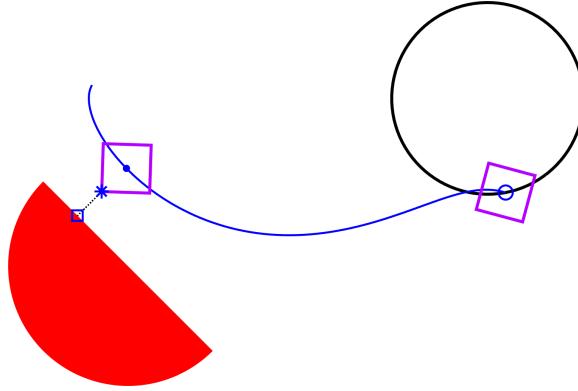
The degree-2 coordinate transformation associated with this orientation is

$$A(s; \varphi) = \begin{bmatrix} \varphi_3 s_1 - \varphi_4 s_2 + \varphi_1 \\ \varphi_3 s_1 + \varphi_4 s_2 + \varphi_2 \end{bmatrix}. \quad (5.38)$$

The shape measure $\mu_s \in \mathcal{M}_+(S \times \Omega)$ is distributed over 6 variables. The size of μ_s 's moment matrix with $k = 2$ at degrees 1-4 is [28, 210, 924, 3003]. The first three distance bounds are $L_2^{1:3} = [2.9158 \times 10^{-5}, 0.059162, 0.14255]$, and MATLAB runs out of memory on the experimental platform at degree 4. A successful recovery is achieved at the degree 3 relaxation, as pictured in Figure 5.15. This rotating-set near-optimal trajectory is encoded by $\varphi_0^* \approx (1.575, -0.3928, 0.2588, 0.9659)$, $t_p^* \approx 3.371$, $s^* \approx (-0.1, 0.1)$, $x_p^* \approx (-0.1096, -0.3998)$, $\varphi_p^* \approx (-0.0064, -0.2921, -0.0322, -0.9995)$, and $y^* \approx (-0.2104, -0.4896)$. Computing this degree-3 relaxation required 75.43 minutes.

5.9 Extensions

This section presents modifications to the distance estimation programs in order to handle systems with uncertainties and distance functions c generated by polyhedral norms.


 Figure 5.15: Rotation, L_2 bound of 0.1425

5.9.1 Uncertainty

Distance estimation can be extended to systems with uncertainty. For the sake of simplicity, this section is restricted to time-dependent uncertainty. Assume that $H \subset \mathbb{R}^{N_h}$ is a compact set of plausible values of uncertainty, and the uncertain process $h(t), \forall t \in [0, T]$ may change arbitrarily in time within H [43]. The distance estimation problem with time-dependent uncertain dynamics is

$$\begin{aligned} P^* = & \inf_{t, x_0, y, h(t)} c(x(t \mid x_0, h(t)), y) \\ & \dot{x}(t) = f(t, x, h(t)), \quad h(t) \in H \quad \forall t \in [0, T] \\ & x_0 \in X_0, \quad y \in X_u. \end{aligned} \quad (5.39)$$

The process $h(t)$ acts as an adversarial optimal control aiming to steer $x(t)$ as close to X_u as possible. The occupation measure μ may be extended to a Young measure (relaxed control) $\mu \in \mathcal{M}_+([0, T] \times X \times H)$ [45, 8].

The Liouville equation (5.11c) may be replaced by $\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx} \mathcal{L}_f^\dagger \mu$, which should be understood to read $\langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle \partial_t v(t, x) + f(t, x, h) \cdot \nabla_x v(t, x), \mu \rangle$ for all test functions $v \in C^1([0, T] \times X)$. Any trajectory with uncertainty process $h(t)$ may be represented by a tuple $(x_0, x_p, t_p, y, h(\cdot))$. This trajectory admits a measure representation similar to the proof of 5.4.1, where the measure μ is the occupation measure of $t \mapsto (t, x(t \mid x_0, h(t)))$ in times $[0, t_p]$. The work in [43] applies a collection of existing uncertainty structures to peak estimation problems (time-independent, time-dependent, switching-type, box-type), and all of these structures may be applied to distance estimation.

To illustrate these ideas, consider the following Flow system with time-dependent uncer-

tainty:

$$\dot{x} = \begin{bmatrix} x_2 \\ (-1 + h)x_1 - x_2 + \frac{1}{3}x_1^3 \end{bmatrix} \quad h \in [-0.25, 0.25]. \quad (5.40)$$

An L_2 distance bound of 0.1691 is computed at the degree 5 relaxation of the uncertain distance estimation program, as visualized in Figure 5.16. The first five distance bounds are $L_2^{1:5} = [5.125 \times 10^{-5}, 1.487 \times 10^{-4}, 0.1609, 0.1688, 0.1691]$.

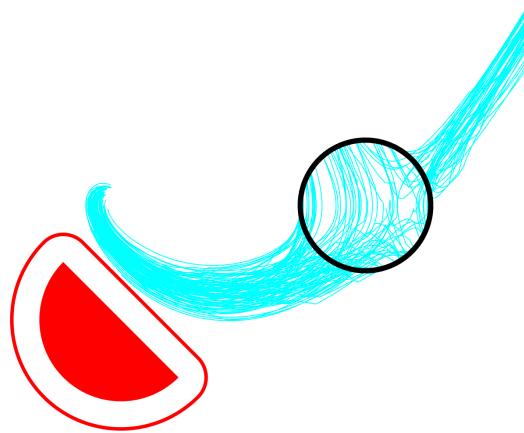


Figure 5.16: Uncertain Flow (5.40), L_2 bound of 0.1691

5.9.2 Polyhedral Norm Penalties

The infinite-dimensional LP (5.11) is valid for all continuous costs $c(x, y) \in C(X^2)$, but its LMI relaxation can only handle polynomial costs $c(x, y) \in \mathbb{R}[x, y]$. The L_p distance is defined as $\|x - y\|_p = \sqrt[p]{\sum_i |x_i - y_i|^p}$ when p is finite and $\|x - y\|_\infty = \max_i |x_i - y_i|$ for p infinite. The cost $\|x - y\|_p^p$ is polynomial when p is finite and even; otherwise the L_p distance has a piecewise definition in terms of absolute values. The theory of convex (LP) lifts may be used to interpret piecewise constraints into valid LMIs [70, 71]. Slack variables $q \in \mathbb{R}$ (or $q_i \in \mathbb{R}$ as appropriate) may be added to form enriched infinite-dimensional LPs. The objective $\langle c, \eta \rangle$ from (5.11a) could be

replaced by the following terms for the examples of L_∞ , L_1 , and L_3 distances:

$$\begin{aligned} \|x - y\|_\infty & \quad \min \quad q \\ -q \leq \langle x_i - y_i, \eta \rangle \leq q & \quad \forall i = 1..n \end{aligned} \tag{5.41a}$$

$$\begin{aligned} \|x - y\|_1 & \quad \min \quad \sum_i q_i \\ -q_i \leq \langle x_i - y_i, \eta \rangle \leq q_i & \quad \forall i = 1..n \end{aligned} \tag{5.41b}$$

$$\begin{aligned} \|x - y\|_3^3 & \quad \min \quad \sum_i q_i \\ -q_i \leq \langle (x_i - y_i)^3, \eta \rangle \leq q_i & \quad \forall i = 1..n. \end{aligned} \tag{5.41c}$$

Distances induced by polyhedral norms can be included through this lifting framework [72]. Figure 5.17 visualizes the near-optimal trajectory for a minimum L_1 distance bound of 0.4003 (cost (5.41c)) at degree 4. This trajectory starts at $x_0^* \approx (1.489, -0.3998)$ and reaches the closest approach between $x_p^* \approx (0, -0.2997)$ and $y^* \approx (-0.1777, -0.5223)$ at time $t^* \approx 0.6181$ units. The first five L_1 distance bounds are $L_1^{1:5} = [3.179 \times 10^{-9}, 4.389 \times 10^{-8}, 0.3146, 0.4003, 0.4003]$.

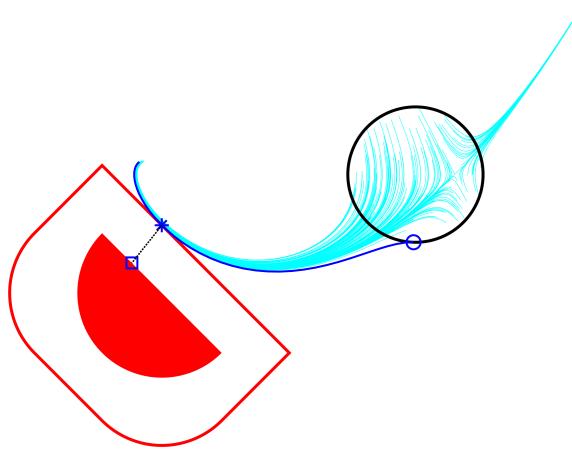


Figure 5.17: Flow system L_1 bound of 0.4003

5.10 Conclusion

This chapter presented an infinite-dimensional linear program in occupation measures to approximate the distance estimation problem. The LP objective is equal to the distance of closest

CHAPTER 5. PEAK ESTIMATION FOR SAFETY ANALYSIS

approach between points along trajectories and points on the unsafe set under mild compactness and regularity conditions. Finite-dimensional truncations of this LP yield a converging sequence of SDP lower bounds to the minimal distance under further conditions (Archimedean). The distance estimation problem can be modified to accommodate dynamics with uncertainty, piecewise distance functions, and movement of shapes along trajectories. Future work includes formulating and implementing control policies to maximize the distance of closest approach to the unsafe set while still reaching a terminal set within a specified time (preliminary work on this subject is detailed in Chapter 8).

Part 2: Robust Safety and Peak Minimizing Control

Chapter 6

Robust Counterparts and Data Driven Analysis

6.1 Background

This chapter analyzes the following disturbance-affine dynamical system:

$$\dot{x}(t) = f(t, x(t), w(t)) = f_0(t, x) + \sum_{\ell=1}^L w_\ell(t) f_\ell(t, x). \quad (6.1)$$

The state $x \in X \subset \mathbb{R}^n$ and the input $w(t) \in W \subset \mathbb{R}^L$ are assumed to lie in compact sets. The time horizon $t \in [0, T]$ is finite for convergence purposes. It is further required that the set W is an L -dimensional compact Semidefinite Representable (SDR) convex set with non-empty interior [73].

An SDR set could arise from a sequence of observations of $\dot{x}(t)$ as corrupted by bounded noise. An example of such an SDR set W is the L -dimensional polytope described by m constraints (up to m faces), which may be expressed as

$$W = \{w \mid Aw \leq b\} \quad A \in \mathbb{R}^{m \times L}, \quad b \in \mathbb{R}^m. \quad (6.2)$$

Letting $x_0 \in X_0$ be an initial condition and $w(t)$ be an admissible control with $w(t) \in W \forall t \in [0, T]$ with T finite, the state obtained by following dynamics in (6.1) is

$$x(t) = x(t \mid x_0, w(\cdot)). \quad (6.3)$$

There is no imposition of continuity on the process $w(t)$.

CHAPTER 6. ROBUST COUNTERPARTS AND DATA DRIVEN ANALYSIS

The problem instances that will be addressed in this chapter are peak estimation, distance estimation, reachable set estimation, and Region of Attraction (ROA) maximization. Each problem instance may be cast as an infinite-dimensional LP and approximated through the moment-SOS hierarchy. Each problem has a Lie derivative nonpositivity constraint that usually induces the largest PSD matrix by numerical solvers. Such a constraint may be split using infinite-dimensional robust counterparts [74] into smaller PSD matrix constraints using convex duality [75] and a theorem of alternatives [76]. Decomposition of SDR sets W move beyond the previously considered box cases in [77] [78] and polytope cases in [69, 79].

Peak estimation finds an initial condition x_0 and input w that maximizes the instantaneous value of a state function $p(x(t))$ along a trajectory [5]. Distance estimation is a variation of peak estimation that finds the distance of closest approach between points along trajectories $x(t \mid x_0, w)$ and an unsafe set X_u (Chapter 5, [58, 13]). Reachable set estimation identifies the set of points X_T such that there exists a pair x_0, w where $x(T \mid x_0, w) \in X_T$ [80]. Peak and reachable set estimation under input-affine and SDR constraints may arise from the data-driven setting where state-derivative observations $\mathcal{D} = \{(t_k, x_k, y_k)\}_{k=1}^{N_s}$ are available, subject to an semidefinite-bounded noise process η ($y(t_k) \doteqdot \dot{x}(t_k) + \eta_k$). An L_∞ -bounded noise model arises from propagating errors from finite-difference schemes to estimate \dot{x} , while an L_2 -bounded noise model is derived from stochastic/chance constraints when \dot{x} has a Gaussian distribution.

Reachable set estimation using LPs occurs from outside in [21] and from inside in [81]. SDPs associated with the moment-SOS hierarchy will produce polynomial sublevel sets that converge in volume to the true reachable set as the polynomial degree increases (under mild conditions, and outside a set of measure zero). Controllers may be formulated to maximize the Backward Reachable Set (BRS), in which the volume of the set of initial conditions X_0 that can be steered towards a target set X_T is maximized [82]. Other approaches towards reachable set estimation of nonlinear systems includes ellipsoidal methods [83], polytopes [84], and interval methods using mixed monotonicity [85]. Infinite-dimensional LPs have also been applied to region of attraction estimation and to backwards reachable set maximizing control [21].

Table 6.1 lists the PSD constraint of maximal size involved in a peak estimation problem (input-affine, polytope) with the WSOS degree $2d = 8$, number of polytope-constrained inputs $L = 10$, state dimension $n = 2$, and dynamics degree $\deg(f) = 3$ (Lie nonpositivity constraint in Section 6.7.3).

The polytope W in this case has 33 faces and 7534 vertices. Performing a size-8568 PSD constraint in solvers such as Mosek or Sedumi is intractable. Applying a vertex decomposition

Table 6.1: Size of largest Lie constraint Gram Matrix (Peak Estimation)

$$\begin{array}{ll} \text{Chapter 4 (4.11)} & \binom{1+n+L+d+\lceil \deg(f)/2 \rceil - 1}{1+n+L} = 8568 \\ \text{This chapter} & \binom{1+n+d+\max_{\ell} \lceil \deg(f_{\ell})/2 \rceil - 1}{1+n} = 56 \end{array}$$

requires that 7534 PSD constraints of size 56 hold. In contrast, a facial decomposition (particular form of the Lie robust counterpart) introduced in this chapter needs only $33 + 1$ PSD constraints of size 56.

Section 6.2 reviews the definition of SDR sets, the concept of robust counterparts, and the notion of Polynomial Matrix Inequalities (PMIs). Section 6.3 poses the peak estimation, distance estimation, reachable set estimation, and BRS maximization programs for systems of the form in (6.1) with convex-bounded uncertainty w . Section 6.4 splits the Lie derivative constraint over the SDR uncertainty W through the use of a facial decomposition. Section 6.5 details polynomial approximation and SOS programs of the robust counterparts. Section 6.6 reviews background of the polyhedral structure of consistency constraints induced by model structures and L_{∞} -bounded noise processes. Section 6.7 presents examples of robust counterparts acting on all four problems. The chapter is concluded in section 6.8. Appendix A.2 contains a proof showing that multiplier functions associated with a certificate of Lie constraint nonnegativity may be chosen to be continuous. Appendix A.3 applies robust counterparts to Lie constraints that possess integral terms (running costs). Appendix A.4 extends a proof from [6] that the auxiliary function for peak estimation may be approximated by a polynomial. Appendix A.5 shows that the multiplier functions may be approximated by polynomials. Appendix A.6 discusses dual measure LPs to the developed infinite-dimensional robust counterparts and applies a moment-based scheme to recover polynomial controller laws.

The polytopic uncertainty work in this chapter is from [69, 79] and was coauthored by Mario Sznaier.

6.2 Preliminaries

This section presents robust counterparts for a linear inequality and PMIs (SOS-matrices) for nonnegativity proofs.

6.2.1 Robust Counterpart

A cone $K \subset \mathbb{R}^n$ is a set such that $\forall c > 0, x \in K \implies cx \in K$ [75]. A cone K defines a partial ordering \geq_K as $x_1 \geq_K x_2$ if $x_1 - x_2 \in K$. The dual K^* of the (finite-dimensional) cone K is the set $\{y \in \mathbb{R}^n \mid x^T y \geq 0 \ \forall x \in K\}$. The cone K is pointed if $x \in K$ and $-x \in K$ implies that $x = 0$.

Definition 6.2.1 ([73]). *Let $S \subset \mathbb{R}^m$ be a set and let $K \subseteq \mathbb{R}^n$ be a cone. The set S is K -representable if there exists a finite dimension q and matrices $A \in \mathbb{R}^{n \times m}, G \in \mathbb{R}^{n \times q}, e \in \mathbb{R}^n$ such that*

$$S = \{x \in \mathbb{R}^m \mid \exists \lambda \in \mathbb{R}^q : Ax + G\lambda + e \in K\}. \quad (6.4)$$

The set S is SDR if K is a subset of the PSD cone (where PSD cone may be vectorized as in $[a, b; b, c] \in \mathbb{S}_+^2 \rightarrow (a, b, c) \in K$).

SDR sets are also referred to as ‘projections of spectahedra’ or ‘spectahedral shadows’. SDR sets form a strict subset of all convex sets. The product, intersection, and projections of SDR sets are all SDR.

This chapter will focus on three specific self-dual cones:

1. Nonnegative ($\mathbb{R}_{\geq 0}$)
2. Second-Order Cone (SOC)/Lorentz ($Q^n : \{(u, v) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \mid \|u\|_2 \leq v\}$)
3. Positive Semidefinite (\mathbb{S}_+^n)

Define the constraint matrices $a_0, a_\ell \in \mathbb{R}^r$ and vectors $b_0, b_\ell \in \mathbb{R}$ for all $\ell = 1..L$. Define W as the intersection of N_s SDR sets with cones $K_1..K_s$ as

$$W = \{w \in \mathbb{R}^L : \forall s = 1..N_s : \exists \lambda_s \in \mathbb{R}^{q_s} : A_s w + G_s \lambda_s + e_s \in K_s\}. \quad (6.5)$$

The following systems are a robust semi-infinite linear inequality constraint in $\beta \in \mathbb{R}^r$ that must hold for all uncertain values w in an SDR:

$$\text{Non-strict : } \forall w \in W : a_0^T \beta + \sum_{\ell=1}^L w_\ell a_\ell^T \beta \leq b_0 + \sum_{\ell=1}^L w_\ell b_\ell \quad (6.6)$$

$$\text{Strict : } \forall w \in W : a_0^T \beta + \sum_{\ell=1}^L w_\ell a_\ell^T \beta < b_0 + \sum_{\ell=1}^L w_\ell b_\ell. \quad (6.7)$$

Definition 6.2.2 ([74] (1.3.14)). *The robust counterpart of (6.6) with respect $w \in W$ is the conic set of constraints in variables $\{\zeta_s\}_{s=1}^{N_s}$*

$$\sum_{s=1}^{N_s} e_s^T \zeta_s + a_0^T \beta \leq b_0 \quad (6.8a)$$

$$G_s^T \zeta_s = 0 \quad \forall s = 1..N_s \quad (6.8b)$$

$$\sum_{s=1}^{N_s} (A_s^T \zeta_s)_\ell + a_\ell^T \beta = b_\ell \quad \forall \ell = 1..L \quad (6.8c)$$

$$\zeta_s \in K_s^* \quad \forall s = 1..N_s. \quad (6.8d)$$

Theorem 6.2.1 (Theorem 1.3.4 of [74]). *Assume that each K_s is a convex and pointed cone with nonempty interior. Further assume that there exists a Slater point $(w_s, \lambda_s : A_s \bar{w}_s + G_s \bar{\lambda}_s + e_s \in \text{int}(K_s))$ for each non-polyhedral cone K_s . Then the semi-infinite program (6.5) is feasible iff the finite-dimensional robust counterpart (6.8) is feasible. Additionally, (6.5) is infeasible iff (6.8) is infeasible.*

Lemma 6.2.2. *Feasibility equivalence of the robust counterpart also holds in the strict case (6.7) by applying a < comparator to (6.8a) [86].*

6.2.2 Polynomial Matrix Inequalities

The symbol $\mathbb{S}^n[x]$ will refer to the set of $n \times n$ symmetric-matrix-valued polynomials in an indeterminate x . The duality paring between two symmetric matrices $A, B \in \mathbb{S}^n$ is $\text{Tr}(AB) = \sum_{ij} A_{ij} B_{ij}$.

The matrix $P \in \mathbb{S}^n[x]$ is an SOS-matrix if there exists a size $s \in \mathbb{N}$, a Gram matrix $Q \in \mathbb{S}_+^{sn}$, and a polynomial vector $v(x) \in \mathbb{R}[x]^s$, such that (Lemma 1 of [87])

$$P(x) = (v(x) \otimes I_s)^T Q (v(x) \otimes I_s). \quad (6.9)$$

The cone of SOS matrices of size n is $\Sigma^n[x] \subset \mathbb{S}^n[x]$, and its degree- $2d$ truncation is $\Sigma_{\leq 2d}^n \subset \Sigma^n[x]$. The BSA set \mathbb{K} in this chapter will be expressed as the locus of PSD polynomial matrix constraints in matrix constraint terms $G_i(x) \in \mathbb{S}^{n_i}(x)$:

$$\mathbb{K} = \{x \in \mathbb{R}^n \mid G_i(x) \succeq 0 \ \forall i = 1..N_c\}. \quad (6.10)$$

Let $q(x) \in \mathbb{R}[x]$ be a polynomial. A PMI over the scalar q with respect to the region \mathbb{K} is

$$q(x) \geq 0 \quad \forall x \in \mathbb{K}. \quad (6.11)$$

The Scherer Psatz proving that $q(x) > 0$ over \mathbb{K} is the statement that (Theorem 2 of [87])

$$q(x) = \sigma_0(x) + \sum_{i=1}^n \text{Tr}(G_i(x)\sigma_i(x)) + \epsilon \quad (6.12a)$$

$$\sigma_0 \in \Sigma[x], \forall i \in 1..N_c : \sigma_i \in \Sigma^{n_i}[x], \epsilon > 0. \quad (6.12b)$$

The WSOS cone $\Sigma[\mathbb{K}]$ is the cone of all polynomials that admit a representation in (6.12) (for $\epsilon \geq 0$). Note that the Scherer Psatz in (6.12) is equivalent to the Putinar Psatz when each constraint term G_i has size $n_i = 1$, $\forall i = 1..N_c$. The set \mathbb{K} in (6.10) is Archimedean if there exists an $R > 0$ such that $R - \|x\|_2^2$ has a Scherer Psatz (6.12) expression. Just as in the Putinar Psatz, the Scherer Psatz describes all positive polynomials over \mathbb{K} when \mathbb{K} is Archimedean (Theorem 2 of [87]).

Define $n_0 = 1$, $d_0 = 0$ for the multiplier σ_0 , and define $d_i = \lfloor \deg G_i / 2 \rfloor$ for the constraints $i = 1..N_c$. The Gram matrices in (6.12b) have size $n_i \binom{n+d-d_i}{d-d_i}$. This Gram size should be compared against the scalarization constraint $\forall y \in \mathbb{R}^{n_i} : y_i^T G_i(x) y_i \geq 0$ involving $n_i + n$ variables, thus resulting in a combinatorially larger Gram matrix of size $\binom{n+n_i+d-d_i}{d-d_i}$.

Refer to [87, 88] for generalizations of the presented Scherer Psatz in (6.12), such as cases where $q(x)$ is a polynomial matrix ($q \in \mathbb{S}^n[x]$ for $n > 1$) over the set \mathbb{K} .

6.3 Analysis and Control Problems

This subsection will present the peak estimation, distance estimation, reachable set estimation, and ROA maximization problems along with their auxiliary function-based approximation approaches. The following assumptions will be shared among all problems,

A1 There is a finite time horizon T .

A2 The state set X is compact with $X_0 \subset X$.

A3 Dynamics f are disturbance-affine (6.1), and all functions $f_0(t, x)$ and $\{f_\ell(t, x)\}_{\ell=1}^L$ are Lipschitz within $[0, T] \times X$.

A4 The input SDR set W satisfies the assumptions of Theorem 6.2.1 (compact, nonempty relative interior, Slater for non-polyhedral cones K_s).

Remark 6.3.1. *The system (6.1) lacks finite escape time by assumptions A1 and A2.*

The Lie derivative $\mathcal{L}_f v(t, x)$ of a scalar *auxiliary* function $v(t, x) \in C^1([0, T] \times X)$ with respect to dynamics in (6.1) is

$$\mathcal{L}_f v = \partial_t v(t, x) + \nabla_x v(t, x) \cdot f(t, x, w) \quad (6.13a)$$

$$= \mathcal{L}_{f_0} v(t, x, w) + \sum_{\ell=1}^L \nabla_x v(t, x) \cdot f_\ell(t, x, w). \quad (6.13b)$$

6.3.0.1 Peak Estimation

The peak estimation problem identifies the maximum value of a state function $p(x)$ attained along trajectories

$$\begin{aligned} P^* &= \sup_{t^* \in [0, T], x_0 \in X_0, w(\cdot)} p(x(t^* | x_0, w(\cdot))) \\ \dot{x}(t) &= f(t, x(t), w(t)), \quad w(t) \in W \quad \forall t \in [0, T]. \end{aligned} \quad (6.14)$$

Assumptions on the cost $p(x)$ are added:

A5 The cost $p(x)$ is continuous inside X .

A6 Some optimal trajectory with $P^* = p(x(t^* | x_0^*, w^*(t)))$ and $t^* \in [0, T]$, $x_0^* \in X_0$, stays in the valid set as $x(t' | x_0^*) \in X$, $w^*(t') \in W$ for all $t' \in [0, t^*]$.

Remark 6.3.2. Further, assumption A5 implies that p is bounded inside X , and therefore that P^* is bounded above.

Peak estimation under polytopic uncertainty is demonstrated in Figure 6.1, with input-affine Flow system dynamics (modified from [36]) of

$$\dot{x}(t) = [x_2(t); -x_1(t) - x_2(t) + (1 + w(t))x_1^3(t)/3] \quad (6.15)$$

$$w(t') \in [-0.5, 0.5] \quad \forall t' \in [0, 5].$$

Figure 6.1 minimizes the vertical coordinate x_2 (red line) for trajectories (cyan curves) starting in $X_0 = \{x \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$ (black circle) and evolves according to Flow dynamics (6.15) for $T = 5$ time units.

An infinite-dimensional LP for peak estimation with variables $v(t, x) \in C^1([0, T] \times X)$, $\gamma \in \mathbb{R}$ under a time-varying disturbance process $w(t) \in W$ is (from (4.8) with no switching

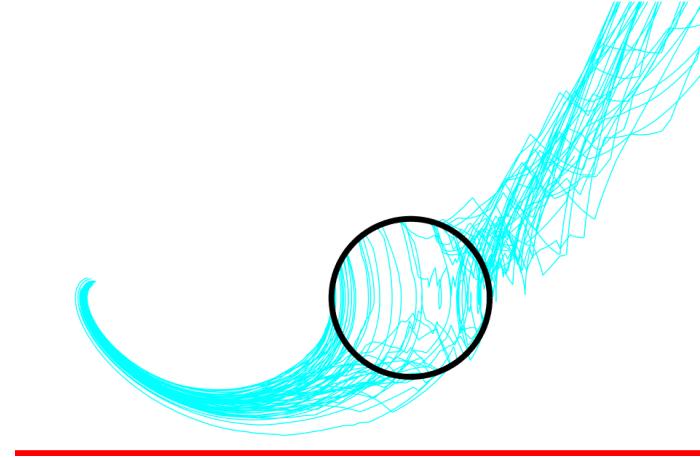


Figure 6.1: Plot of Uncertain Flow system (6.15) trajectories

and with $\Theta = 0$)

$$d^* = \inf_{v,\gamma} \gamma \quad (6.16a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (6.16b)$$

$$\mathcal{L}_f v(t, x, w) \leq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W \quad (6.16c)$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (6.16d)$$

$$v(t, x) \in C^1([0, T] \times X). \quad (6.16e)$$

The auxiliary function $v(t, x)$ is an upper bound on the cost $p(x)$ (6.16d), and must decrease along all possible disturbed trajectories (6.16c). The $P^* = d^*$ between programs (6.14) and (6.16) will match under assumptions A1-A6. The LP in (6.16) may be approximated through the moment-SOS hierarchy as reviewed in Section 2.4, and this sequence of upper bounds (outer approximations) will converge $d_d^* \geq d_{d+1}^* \geq \dots$ to P^* .

The order-4 SOS peak estimate of the Flow system scenario in (6.15) starting in X_0 yields $x_2(t) \geq -0.7862$, as shown in the red line in Figure 6.1.

6.3.0.2 Distance Estimation

Distance estimation from Chapter 5 will be solved as a peak estimation program over $p(x) = -c(x; X_u)$ for a compact unsafe set X_u . The Distance Estimation program under uncertainty

$w(\cdot)$ (with a sign difference from (5.17)) is

$$d^* = \inf_{v(t,x), \gamma} \gamma \quad (6.17a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (6.17b)$$

$$\mathcal{L}_f v(t, x, w) \leq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W \quad (6.17c)$$

$$v(t, x) \geq \phi(x) \quad \forall (t, x) \in [0, T] \times X \quad (6.17d)$$

$$\phi(x) \geq c(x, y) \quad \forall (x, y) \in X \times Y. \quad (6.17e)$$

$$\phi \in C(X), v \in C^1([0, T] \times X). \quad (6.17f)$$

The distance of closest approach is $c^* = -d^*$.

6.3.0.3 Reachable Set Estimation

The reachability set X_T is the set of all x that can be reached at time index $t = T$ for trajectories starting in the set X_0 (under assumptions A1-A4):

$$X_T = \{x(T \mid x_0) \mid x(0) = x_0 \in X_0, x'(t) = f(t, x, w)\}. \quad (6.18)$$

The methods in [21] propose the following volume maximization problem to find the reachable set X_T by

$$P^* = \sup_{X_T \subset X} \text{vol}(X_T) \quad (6.19a)$$

$$\forall \tilde{x} \in X_T, \exists x_0 \in X_0, w(t) \in W : \quad (6.19b)$$

$$\tilde{x} = x(T \mid x_0, w(t)) \quad (6.19b)$$

$$x'(t) = f(t, x) \quad \forall t \in [0, T]. \quad (6.19c)$$

An infinite-dimensional LP in continuous functions $v(t, x)$ and $w(x)$ may be developed to

outer-approximate the reachable set X_T [21] as

$$d^* = \inf_{v(t,x), \phi(x)} \int_X \phi(x) dx \quad (6.20a)$$

$$v(0, x) \leq 0 \quad \forall x \in X_0 \quad (6.20b)$$

$$\phi(x) + v(T, x) \geq 1 \quad \forall x \in X \quad (6.20c)$$

$$\forall (t, x, w) \in [0, T] \times X \times W :$$

$$\mathcal{L}_f v(t, x, w) \leq 0 \quad (6.20d)$$

$$v(t, x) \in C^1([0, T] \times X) \quad (6.20e)$$

$$\phi(x) \in C_+(X). \quad (6.20f)$$

At a degree- d LMI relaxation, the set $\{x \in X \mid \phi(x) \geq 1\}$ is an outer approximation to the reachable set with volume bounds yielding the bounds $d_d^* \geq d_{d+1}^* \geq P^* = \text{vol}(X_T)$. This sublevel set will converge in volume to the region of attractions (excluding sets of measure zero) as $d \rightarrow \infty$. The level set approximations will be valid except for possibly a set with Lebesgue measure zero (e.g., points, planes). Inner approximations to the region of attraction can be performed through the methods in [81].

6.3.0.4 Region of Attraction Maximization

Let $X_T \subsetneq X$ be a given ‘goal’ or ‘target’ set. The BRS/ROA given X_T is the set

$$X_0 = \{x_0 \mid x(0) = x_0 \in X_0, x'(t) = f(t, x, w(t)), \quad (6.21)$$

$$x(T \mid x_0, w) \in X_T, w(t) \in W\}.$$

Intuitively, the set X_0 is the set of states that may be steered towards the goal set X_T in time T . The ROA-maximization formulation of optimal control aims to maximize the volume of X_0 , similar to how problem (6.19) maximized the volume of X_T to acquire the reachable set. The LP in functions

v, ϕ to perform ROA maximization is [21]

$$d^* = \inf \int_X \phi(x) dx \quad (6.22a)$$

$$v(T, x) \geq 0 \quad \forall x \in X_T \quad (6.22b)$$

$$\phi(x) \geq 1 + v(0, x) \quad \forall x \in X \quad (6.22c)$$

$$\forall(t, x, w) \in [0, T] \times X \times W :$$

$$\mathcal{L}_f v(t, x, w) \leq 0 \quad (6.22d)$$

$$v(t, x) \in C^1([0, T] \times X) \quad (6.22e)$$

$$\phi(x) \in C_+(X). \quad (6.22f)$$

Note that the roles of $t = \{0, T\}$ and some signs are swapped in (6.22) as compared to (6.20).

6.4 Decomposed Lie Constraint

The Lie constraints in (6.16c), (6.17c), (6.20d), (6.22d) are the main decomposable expressions to be addressed in this work. Each case requires that an auxiliary function $v(t, x) \in C^1([0, T] \times X)$ be non-increasing along trajectories of f given by (6.1). The Lie derivative in (6.13) must respect the constraint

$$\mathcal{L}_f v(t, x, w) \leq 0 \quad \forall(t, x, w) \in [0, T] \times X \times W. \quad (6.23)$$

Lemma 6.4.1. *Constraint (6.23) may be expressed as a semi-infinite linear inequality (6.6) under the correspondence (holding $\forall \ell = 1..L$)*

$$b_0 = -(\partial_t + f_0(t, x) \cdot \nabla_x) v(t, x) = -\mathcal{L}_{f_0} v(t, x) \quad a_0 = 0 \quad (6.24a)$$

$$b_\ell = -f_\ell(t, x) \cdot \nabla_x v(t, x) \quad a_\ell = 0. \quad (6.24b)$$

Theorem 6.4.2. *The robust counterpart of (6.23) with (possibly discontinuous) multiplier variables $\zeta_s(t, x)$ is*

$$\mathcal{L}_{f_0} v(t, x) + \sum_{s=1}^{N_s} e_s^T \zeta_s(t, x) \leq 0 \quad \forall[0, T] \times X \quad (6.25a)$$

$$G_s^T \zeta_s(t, x) = 0 \quad \forall s = 1..N_s \quad (6.25b)$$

$$\sum_{s=1}^{N_s} (A_s^T \zeta_s(t, x))_\ell + f_\ell(t, x) \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \quad (6.25c)$$

$$\zeta_s(t, x) \in K_s^* \quad \forall s = 1..N_s, (t, x) \in [0, T] \times X. \quad (6.25d)$$

Feasibility of the robust counterpart is equivalent to feasibility (6.23) under Assumptions A1-A4.

Proof. This follows from Theorem 6.2.1 (equivalence of robust counterpart) applied to the correspondence in Lemma 6.4.1. \square

Theorem 6.4.3. *The multiplier functions ζ in (6.25d) may be chosen to be continuous when $\mathcal{L}_f v(t, x, w) < 0$ holds strictly (from Lemma 6.2.2).*

Proof. See Appendix A.2. \square

Appendix A.3 formulates robust counterparts to (6.23) under commonly used integral costs J , producing the constraint $\mathcal{L}_f v(t, x, w) + J(t, x, w) \geq 0$.

6.5 Polynomial Approximation

This section develops polynomial and SOS approximations of the infinite-dimensional Lie robust counterpart (6.25).

Theorem 6.5.1. *Given a tolerance $\epsilon > 0$, the peak estimation task (6.16) admits a polynomial auxiliary function $V(t, x)$ with objective $d^* + (5/2)\epsilon$ such that $\mathcal{L}_f V(t, x) < 0$ holds strictly in $[0, T] \times X$*

Proof. See Appendix A.4. \square

Remark 6.5.1. *Refer to Theorem 9.3.3 for a similar proof w.r.t. distance estimation, and to [21, 77] for proofs of no relaxation gap (in the sense of volume) for reachable set estimation and for BRS maximization.*

Theorem 6.5.2. *Multipliers ζ in (6.25d) can be chosen to be polynomial when v is polynomial and when (6.25a) holds strictly.*

Proof. See Appendix A.5. When v is polynomial, the vector indexed by $b_\ell = f_\ell(t, x) \cdot \nabla_x v(t, x)$ is also polynomial. \square

We now provide details on polynomial approximation and SOS implementation over the nonnegative, SOC, and PSD cones. These details can be combined to perform SOS approximation of sets involving multiple cones. It is not required that the set W be Archimedean, only that W must be compact (A1).

6.5.1 Polytope Restriction

Assume that the SDR set W is the polytope $\{w \mid \exists \lambda \in \mathbb{R}^{L'} : Aw + G\lambda \leq e\}$ for matrices $A \in \mathbb{R}^{m \times L}, G \in \mathbb{R}^{m \times L'}, b \in \mathbb{R}^m$. Define A_s as the s -th row of A and $(A^T)_\ell$ as the ℓ -th column of A (transpose ℓ -th row of G^T). This case corresponds to $\forall s = 1..m : K_s = \mathbb{R}_{\geq 0}$ under the cone description $(-A_s, -G_s, e_s)$ (6.5). The expression of the robustified Lie constraint in (6.25) for the polytopic case is

$$\mathcal{L}_{f_0}v(t, x) + e^T \zeta(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (6.26a)$$

$$-(A^T)_\ell \zeta(t, x) + f_\ell \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \quad (6.26b)$$

$$G^T \zeta(t, x) = 0 \quad (6.26c)$$

$$\zeta(t, x) \in \mathbb{R}_{\geq 0}^m. \quad (6.26d)$$

The SOS tightening of the constraints in (6.26a) when (v, ζ) are polynomials is

$$\mathcal{L}_{f_0}v(t, x) + e^T \zeta(t, x) \in \Sigma^1([0, T] \times X) \quad (6.27a)$$

$$\text{coeff}_{t,x}(-A^T \zeta(t, x) + f_\ell \cdot \nabla_x v(t, x)) = 0 \quad (6.27b)$$

$$\text{coeff}_{t,x}(-G^T \zeta(t, x)) = 0 \quad (6.27c)$$

$$\zeta_s(t, x) \in \Sigma([0, T] \times X) \quad \forall s = 1..m. \quad (6.27d)$$

The degree- d tightening of program 6.27 has a Gram matrix of maximal size size $\binom{n+d}{n}$ from (6.27a) and m Gram matrices of maximal size $\binom{n+d}{d}$ from (6.27d).

6.5.2 Semidefinite Restriction

This subsection involves the case where W is an SDR set with describing matrices $A_0, A_\ell, G_k \in \mathbb{S}^q$:

$$W = \{w \in \mathbb{R}^L \mid A_0 + \sum_{\ell=1}^L w_\ell A_\ell + \sum_{k=1}^{L'} \lambda_k G_k \succeq 0\}. \quad (6.28)$$

The robust counterpart expression in (6.25) is

$$\mathcal{L}_{f_0}v(t, x) + \text{Tr}(A_0 \zeta(t, x)) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (6.29a)$$

$$\text{Tr}(A_\ell \zeta(t, x)) + f_\ell \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \quad (6.29b)$$

$$\text{Tr}(G_k \zeta(t, x)) = 0 \quad \forall k = 1..L' \quad (6.29c)$$

$$\zeta(t, x) \in \mathbb{S}_+^q. \quad (6.29d)$$

The Scherer Psatz (6.12) applied to (6.29) is

$$-\mathcal{L}_{f_0}v(t, x) - \text{Tr}(A_0\zeta(t, x)) - \epsilon \in \Sigma^1([0, T] \times X)] \quad (6.30a)$$

$$\text{coeff}_{t,x}(\text{Tr}(A_\ell\zeta(t, x)) + f_\ell \cdot \nabla_x v(t, x)) = 0 \quad \forall \ell = 1..L \quad (6.30b)$$

$$\text{coeff}_{t,x}(\text{Tr}(G_k\zeta(t, x)) = 0) \quad \forall k = 1..L' \quad (6.30c)$$

$$\zeta(t, x) - \epsilon I \in \Sigma^q([0, T] \times X)]. \quad (6.30d)$$

The maximal-size Gram matrix at degree d will either occur in (6.30a) with size $\binom{n+d}{n}$ or in (6.30) with size $q\binom{n+d}{d}$.

6.5.3 Second-Order Cone Restriction

This final subsection involves the SOC case $W = \{w \in \mathbb{R}^L \mid \exists \lambda \in \mathbb{R}^{L'} : \|Aw + G\lambda + e\|_2 \leq r\}$ for $A \in \mathbb{R}^{m \times L}$, $G \in \mathbb{R}^{m \times L'}$, $e \in \mathbb{R}^L$, $r \in \mathbb{R}_{\geq 0}$. The constraint in W may be formulated as the SOC expression

$$(Aw + G\lambda + e, r) \in Q^m. \quad (6.31)$$

The robust counterpart (6.25) applied to (6.31) involves a partitioned multiplier function $\zeta = (\beta, \tau) \in Q^m$:

$$\mathcal{L}_{f_0}v(t, x) + r\tau(t, x) \leq 0 \quad (6.32a)$$

$$G^T\beta(t, x) = 0 \quad (6.32b)$$

$$(A^T)_\ell\beta(t, x) = f_\ell(t, x) \cdot \nabla_x v(t, x) \quad \forall \ell = 1..L \quad (6.32c)$$

$$(\beta(t, x), \tau(t, x)) \in Q^m. \quad (6.32d)$$

An SOS formulation of (6.32) requires the following lemma:

Lemma 6.5.3. *The SOC membership $(\beta, \tau) \in Q^m$ may be expressed by the following SDP with 2×2 blocks [89]*

$$(\beta, \tau) \in Q^m \Leftrightarrow \exists \omega \in \mathbb{R}^m : \begin{bmatrix} \tau & \beta_j \\ \beta_j & \omega_j \end{bmatrix} \in \mathbb{S}_+^2, \quad \tau = \sum_{j=1}^m \omega_j. \quad (6.33)$$

Lemma 6.5.3 will be used to form an SOS-matrix representation of (6.32d):

$$-\mathcal{L}_{f_0}v(t, x) - r\tau(t, x) \in \Sigma^1([0, T] \times X)] \quad (6.34a)$$

$$\text{coeff}_{t,x}(G^T \beta(t, x)) = 0 \quad (6.34b)$$

$$\text{coeff}_{t,x}((A^T)_\ell \beta(t, x) - f_\ell(t, x) \cdot \nabla_x v(t, x)) = 0 \quad \forall \ell = 1..L \quad (6.34c)$$

$$\begin{bmatrix} \sum_{j=1}^m \omega_j(t, x) & \beta_j(t, x) \\ \beta_j(t, x) & \omega_j(t, x) \end{bmatrix} \in \Sigma^2([0, T] \times X]) \quad \forall j = 1..m \quad (6.34d)$$

$$\omega_j(t, x), \beta_j(t, x) \in \mathbb{R}[t, x] \quad \forall j = 1..m. \quad (6.34e)$$

The degree- d truncation of (6.34) involving polynomials $\omega_j, \beta_j \in \mathbb{R}[t, x]_{\leq 2d}$ will have m maximal-size Gram matrices of size $2\binom{n+d}{n}$ from constraint (6.34d).

Remark 6.5.2. The rotated SOC cone is $Q_r^n = \{(u, v, z) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}^2 \mid \|u\|_2 \leq vz\}$ [89]. Membership in Q_r^n may be expressed as a linear transformation of membership in Q^{n+1} by

$$(u, v, z) \in Q_r^n \Leftrightarrow ([u, v - z], v + z) \in Q^{n+1}. \quad (6.35)$$

The identity (6.35) can be used to form SOS-matrix programs from (6.34) for rotated-SOC constrained uncertainty sets W .

6.5.4 Approximation Result

The following theorem summarizes the above restrictions.

Theorem 6.5.4. Assume that $[0, T] \times X$ is Archimedean in addition to A1-A4. Let the SDR cone K from the W -representation (6.5) be the product of nonnegative, SOC, and PSD cones. Then the SOS programs derived from (6.25) (by example (6.27), (6.30), (6.34)) will converge to the strict version of (6.23) when v is polynomial.

Proof. The multiplier functions ζ may be chosen to be polynomial by Theorem 6.5.2. The Archimedean condition of $[0, T] \times X$ ensures that SOS-matrices will generate all positive PSD matrices over $[0, T] \times X$. Because a polynomial ζ exists by 6.5.2, it will be found at some finite-degree SOS tightening, thus proving the theorem. \square

Remark 6.5.3. When (6.23) holds strictly, Theorems 6.5.2 and 6.5.4 can be extended to cases where (A, G, e) in (6.5) are continuous (polynomial) functions of (t, x) in the compact space $[0, T] \times X$.

We illustrate the robust decomposition of the Lie constraint on a peak estimation problem (6.16) under polytopic uncertainty from Section 6.5.1 with

$$d^* = \inf_{v(t,x), \gamma} \gamma \quad (6.36a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (6.36b)$$

$$\mathcal{L}_{f_0} v(t, x) + e^T \zeta(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (6.36c)$$

$$-(A^T)_\ell \zeta(t, x) + f_\ell \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \quad (6.36d)$$

$$G^T \zeta(t, x) = 0 \quad (6.36e)$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (6.36f)$$

$$v(t, x) \in C^1([0, T] \times X) \quad (6.36g)$$

$$\zeta_j(t, x) \in C_+([0, T] \times X) \quad \forall j = 1..m. \quad (6.36h)$$

Appendix A.6 describes the dual problem to (6.36) in terms of occupation measures and explains how approximate control laws may be extracted.

6.6 Data-Driven Setting

This section reviews the L_∞ bounded noise setting and its derived polytopic input constraints for W [90, 91]. We note that other input sets in a data-driven framework include elementwise L_1 noise (sparse channel disturbances), elementwise L_2 noise [92] (e.g., Chi-squared chance constraints on a Gaussian distribution), and semidefinite energy-bounded-noise [93].

Samples y of an unknown continuous-time system $\dot{x} = F(t, x)$ are observed according to the relation $\dot{x}_{observed} = y = F(t, x) + \eta$ with noise term $\|\eta\|_\infty \leq \epsilon_w$. We are also given the knowledge that there exists at least one ground-truth choice of parameters $w^* \in \mathbb{R}^L$ with

$$F(t, x) = f_0(t, x) + \sum_{\ell=1}^L w_\ell^* f_\ell(t, x), \quad (6.37)$$

where the parameters w^* are *a-priori* unknown. The function f_0 represents prior knowledge of system dynamics F , and the dictionary functions $\{f_\ell\}$ serve to describe unknown dynamics.

The tuples $\mathcal{D}_k = (t_k, x_k, y_k)$ for $k = 1..N_s$ observations are contained in the data \mathcal{D} . System parameters w that are consistent with data in \mathcal{D} form a set

$$W = \{w \in \mathbb{R}^L \mid \forall k : \|y_k - f(t_k, x_k; w)\|_\infty \leq \epsilon_w\}. \quad (6.38)$$

CHAPTER 6. ROBUST COUNTERPARTS AND DATA DRIVEN ANALYSIS

The L_∞ term $\|y_k - f(t_k, x_k; w)\|_\infty$ describing W for each record k may be expanded into

$$\|y_k - f_0(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_\ell(t_k, x_k)\|_\infty. \quad (6.39)$$

This L_∞ constraint may be split into n absolute value constraints

$$|y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k)| \leq \epsilon_w. \quad (6.40)$$

The positive and negative side of each absolute value constraint are

$$y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon_w \quad (6.41a)$$

$$y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) \geq -\epsilon_w. \quad (6.41b)$$

By sending all w_ℓ terms to the left-hand side as in

$$-\sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon_w - y_{ik} + f_{i0}(t_k, x_k) \quad (6.42a)$$

$$\sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon_w + y_{ik} - f_{i0}(t_k, x_k), \quad (6.42b)$$

new terms (Γ, h) may be defined as

$$\Gamma_{ik\ell} = [-f_{i\ell}(t_k, x_k); f_{i\ell}(t_k, x_k)] \quad (6.43a)$$

$$h_{ik} = \begin{bmatrix} -y_{ik} + f_{i0}(t_k, x_k) \\ y_{ik} - f_{i0}(t_k, x_k) \end{bmatrix}. \quad (6.43b)$$

The polytope W may be described in terms of constants in (6.43) (equivalent to Equation (6.42)) as

$$W = \left\{ w \in \mathbb{R}^L \mid \forall i, k : \sum_{\ell=1}^L \Gamma_{ik\ell} w_\ell \leq \epsilon_w + h_{ik} \right\}, \quad (6.44)$$

which will be written concisely as the polytope $W = \{w \mid \Gamma w \leq (\epsilon_w + h)\}$.

Remark 6.6.1. *The compactness and non-emptiness assumption of A4 is satisfied when the L_∞ bound ϵ_w is finite and sufficiently many observations in \mathcal{D} are acquired.*

The set W as described in (6.44) has $m = 2N_s n$ affine constraints, most of which are redundant. These redundant constraints can be identified and dropped through the LP method of [94]. The multiplier term ζ is m -dimensional, so lowering m by eliminating redundant constraints is essential in creating tractable problems.

6.7 Examples

Code to execute Lie robust counterparts for analysis and control and to replicate figures and experiments is available at https://github.com/Jarmill/data_driven_occ. All source code was developed in MATLAB 2021a. Dependencies include YALMIP [48] to form the SDPs and MOSEK [49] to solve them. Unless otherwise specified, the SDR uncertainty set W will be polytopic. Redundant constraints (non-exposed faces) in the polytope W were identified and dropped through the LP method of [94]. When W is polytopic, the $w(t)$ inputs of trajectory samples (data-driven analysis) were acquired through hit-and-run sampling of the polytope W [95] as implemented by [96]. In the case of semidefinite-bounded noise, the input $w(t)$ was chosen by choosing a uniformly random direction θ on the $(L - 1)$ -sphere and solving the LMI $\max_{w \in W} \theta^T w$.

6.7.1 SIR System

A demonstration of the polytopic uncertainty arising from the data driven setting may take place on an epidemic example. A basic compartmental epidemic model involves three states: S (susceptible), I (infected), and R (removed). The population is assumed to be normalized such that $S + I + R = 1$. Temporal dynamics of the Susceptible, Infected, Removed (SIR) system with parameters (β, γ) are

$$S' = -\beta SI \quad I' = \beta SI - \gamma I. \quad (6.45)$$

The R trajectory with the dynamics $R' = \gamma I$ may be recovered by the relation $R = 1 - S - I$.

Figure 6.2a plots 100 true and $\epsilon = 0.1$ -corrupted observations of the SIR system with a ground truth of $\beta = 0.4$, $\gamma = 0.1$. Each data-record enforces 4 constraints (positive and negative sides for S and I), and there are 400 affine constraints in total. The 5-sided polytope Θ_D is plotted in Figure 6.2b along with its Chebyshev center in the asterisk at $(\beta_{\text{cheb}}, \gamma_{\text{cheb}}) = (0.0977, 0.4003)$ (the Chebyshev center of a polytope is the center of the inscribed sphere with maximum radius).

Only 5 out of the 400 constraints describing Θ_D in this SIR example are non-redundant. The active constraints in Figure 6.3 are dotted black lines, all other inactive constraints are the gray dotted lines. The polytope Θ_D observed in Figure 6.2b is bordered by solid black lines.

Peak estimation to bound the maximum value of the infected population I is performed on the (6.45) with observations in Fig. 6.2a. The $L = 2$ uncertain parameters are (β, γ) . The (β, γ) consistency set aligning with the observed data is the 5-sided polytope in Fig. 6.2. The peak estimate over a time horizon of $T = 40$ days is $I_{\max} \leq 0.511$ at an order-3 SOS tightening of (6.36).

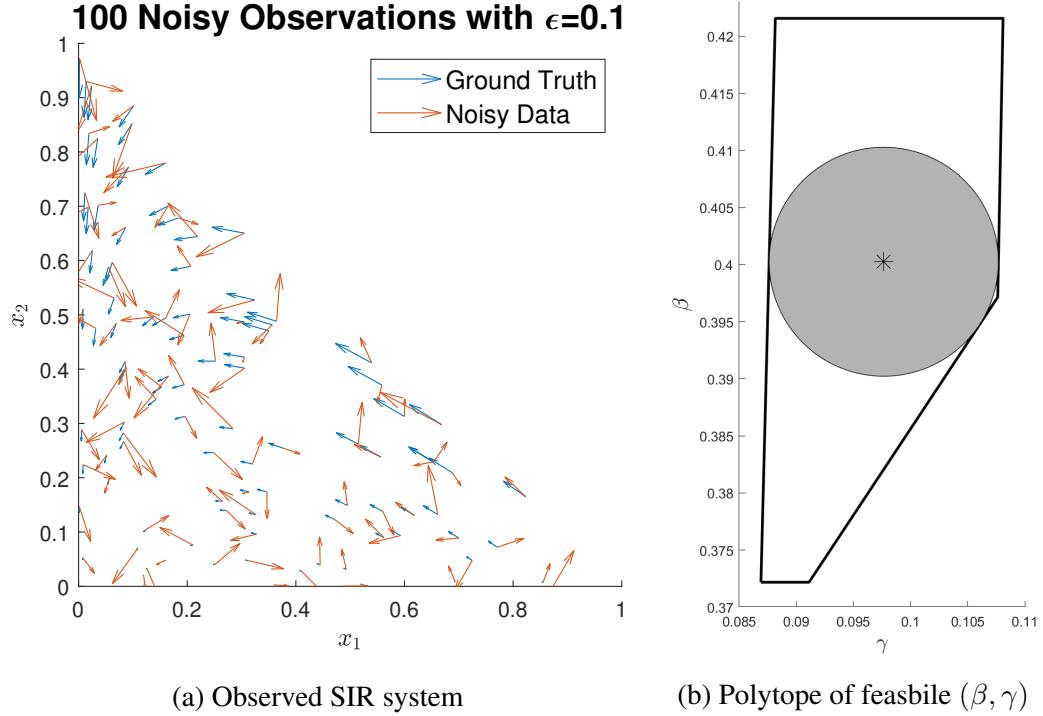


Figure 6.2: Observed and Corrupted SIR

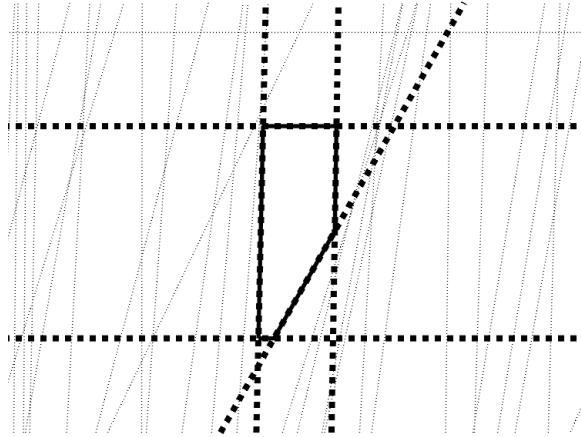
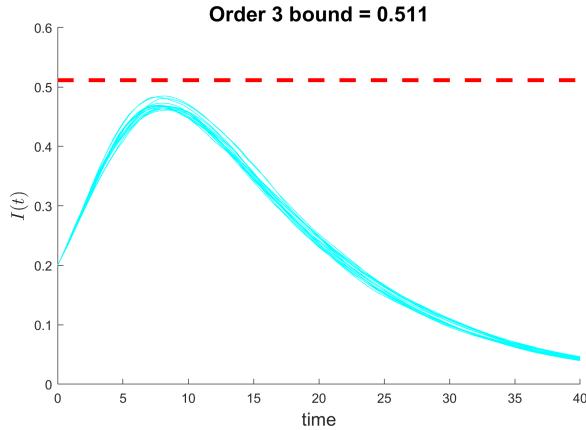
6.7.2 Semidefinite-Disturbed Flow System

This subsection performs peak estimation of the Flow system (6.15) under a semidefinite-constrained disturbance process:

$$f(t, x, w) = \begin{bmatrix} x_2 \\ -x_1 - x_2 + x_1^3/3 + w_1 x_1 + w_2 x_1 x_2 + w_3 x_3 \end{bmatrix} \quad (6.46a)$$

$$W = \left\{ w \in \mathbb{R}^3 : \begin{bmatrix} 1 & w_1 & w_2 \\ w_1 & 1 & w_3 \\ w_2 & w_3 & 1 \end{bmatrix} \in \mathbb{S}_+^3 \right\}. \quad (6.46b)$$

The set in (6.46b) is the standard convex ellotope/pillow spectahedron. Dynamics in (6.46) start at $X_0 = [1.25; 0]$ and continue for $T = 5$ time units in the space $X = [-0.5, 1.75] \times [-1, 0.5]$. The first 6 bounds of maximizing $p(x) = -x_2$ along these trajectories, after performing a robust counterpart, are $p_{1:6}^* = [1, 1, 0.8952, 0.8477, 0.8471, 0.8470]$. At order 6, the largest PSD matrix constraint (for the 3×3 SOS-matrix) is of size $3 \binom{3+6}{6} = 252$ in the variables (t, x) after performing robust decomposition, while the non-decomposed largest PSD size is $3 \binom{6+6}{6} = 2772$ in the variables


 Figure 6.3: Active and inactive constraints describing Θ_D in Figure 6.2b

 Figure 6.4: Maximum value of $I(t)$ over a time horizon of $T = 40$, with unknown (β, γ)
 $(t, x, w).$

6.7.3 Data-Driven Flow System

This subsection will further extend the disturbed Flow system in (6.15) by the case where \dot{x}_2 is modeled by a cubic polynomial $\dot{x}_2 = \sum_{\deg \alpha \leq 3} w_\alpha x_1^{\alpha_1} x_2^{\alpha_2}$ with 10 unknown parameters/inputs $\{w_\alpha\}$. The derivative $\dot{x}_1 = x_2$ remains known and there is no uncertainty in the first coordinate. Table 6.1 in the Introduction refers to this cubic Flow setting.

Figure 6.6 visualizes $N = 40$ observed data points sampled within the initial set $X_{sample} = \{x \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$. The true derivative values are the blue arrows and the $\epsilon = [0; 0.5]$ -corrupted derivative observations are orange. The $N = 40$ points yield $2N = 80$ affine constraints,

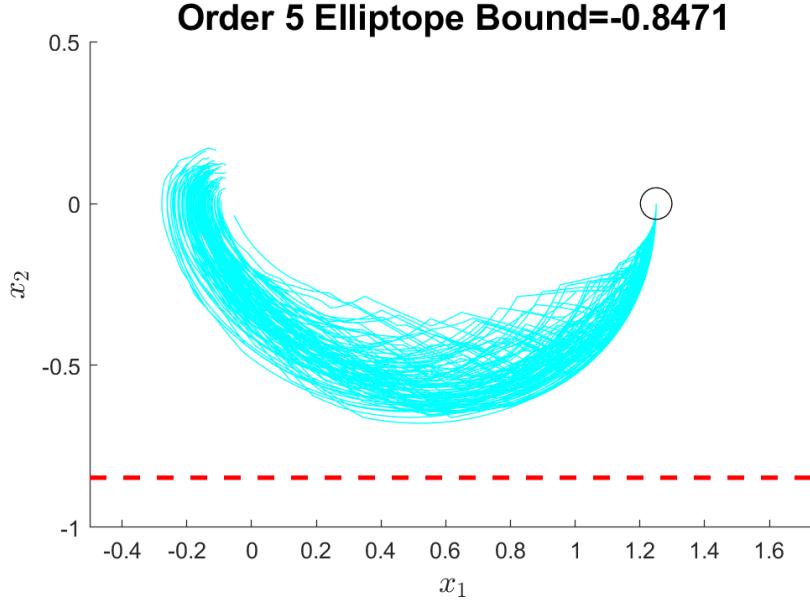


Figure 6.5: Order-5 bound on minimal x_2 for Flow (6.46) under elliptope-constrained noise

of which the polytope W has $L = 33$ faces (nonredundant constraints) and 7534 vertices.

Figure 6.7 displays Flow system trajectories for a time horizon of $T = 5$ starting from the point $X_0 = (1.5, 0)$ (left, Figure 6.7a) and from the circle $X_0 = X_{\text{sample}}$ (right, Figure 6.7b). Each case desires to maximize $p(x) = -x_2$ over the state region of $X = \{x \mid \|x\|_2^2 \leq 8\}$. The first 4 SOS peak estimates in the point X_0 case (Figure 6.7a) are $d_{1:4}^* = [2.828, 2.448, 1.018, 0.8407]$. The first four estimates in the disc X_0 case (Figure 6.7b) are $d_{1:4}^* = [2.828, 2.557, 1.245, 0.894]$.

Figure 6.8 displays the result of distance estimation on the Flow system (3.9) with $N = 40$ points and $\epsilon = [0; 0.5]$. The dynamics model is the same as in the peak (6.7) with $\dot{x} = [x_2; \text{cubic}(x_1, x_2)]$. The initial point is $X_0 = [1; 0]$ in the state set $X = [-1, 1.25] \times [-1.25, 0.7]$, and trajectories are tracked for $T = 5$ time units. The distance function is the L_2 distance and the red half-circle unsafe set is $X_u = \{x \mid 0.5^2 \geq -(x_1 + 0.25)^2 - (x_2 + 0.7)^2, -(x_1 + 0.25)/\sqrt{2} + (x_2 + 0.7)^2/\sqrt{2} \geq 0\}$. The first 5 bounds of the robust distance estimation program are $c_{1:5}^* = [1.698 \times 10^{-5}, 0.1936, 0.2003, 0.2009, 0.2013]$.

6.7.4 Twist System

This section performs peak estimation on the Twist system in (5.34)

A total of $N = 100$ observations with a noise bound of $\epsilon = 0.5$ are taken, and are plotted in Figure 6.9. These $N = 100$ observations will induce $2Nn = 600$ affine constraints on eventual

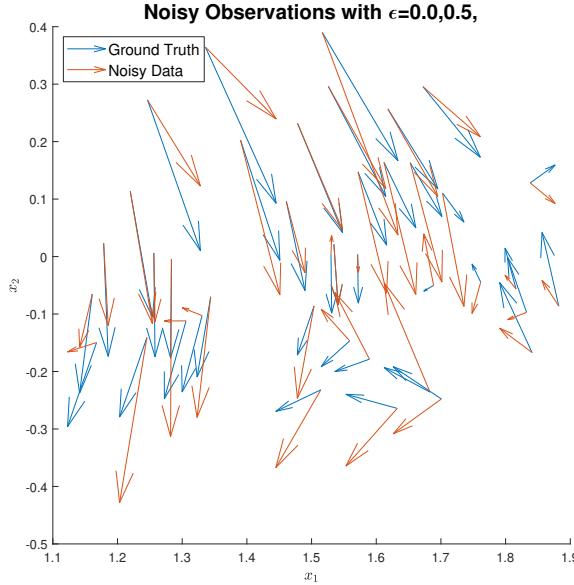


Figure 6.6: Observed data of Flow system (3.9) within a circle

polytopes W .

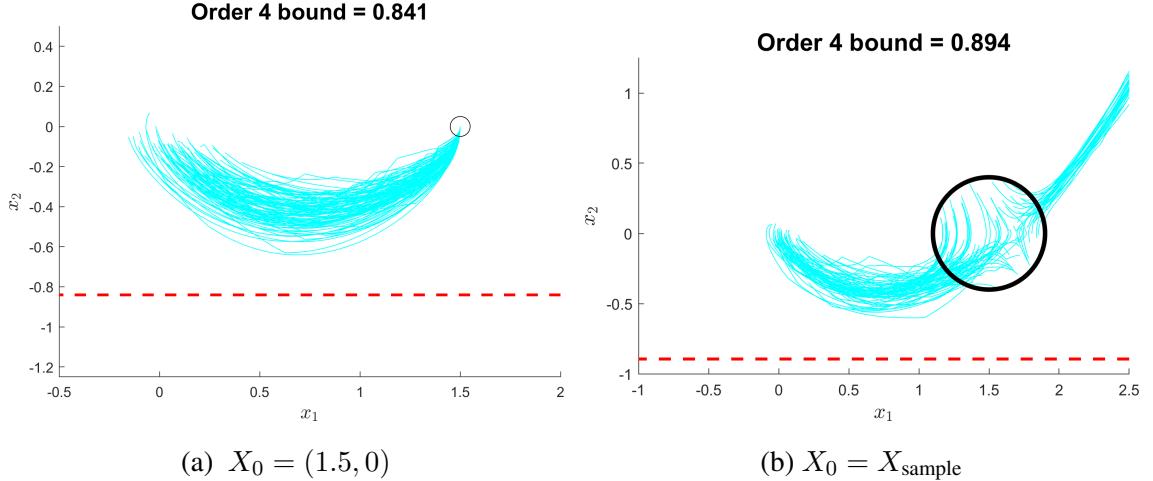
All scenarios in this subsection will find peak estimates on the maximum value of $p(x) = x_3$ of the Twist system over the space $X = \{x \mid -1 \leq x_1, x_2 \leq 1, 0 \leq x_3 \leq 1\}$ and time horizon $T = 8$ starting at $X_0 = [-1, 0, 0]$.

Figure 6.10 involves the $L = 9$ case where B^1 is unknown (left) or when B^3 is unknown (right). The unknown B^1 case in Figure 6.10a has a polytope W with $m = 30$ faces and peak bounds of $d_{1:3}^* = [1.000, 0.9050, 0.8174]$. The known B^3 case in Figure 6.10b also has $m = 30$ faces in its polytope with peak bounds of $d_{1:3}^* = [1.000, 0.9050, 0.8174]$. The maximal PSD matrix size of the Lie nonpositivity constraint is 2380 pre-decomposition and 70 post-decomposition.

When both parameters (B^1, B^3) are unknown, the polytope W has $L = 18$ dimensions and $m = 70$ nonredundant faces. The first peak estimates on this system are $d_{1:2}^* = [1.000, 0.9703]$ as plotted in Figure 6.11. At degree 2, the maximal PSD matrix size in the Lie constraint falls from 2300 pre-decomposition and to 35 post-decomposition. The experimental platform became unresponsive in YALMIP when attempting to compile the degree $d = 3$ model.

6.7.5 Reachable Set Example

Figure 6.12 illustrates data-driven reachable set estimation on the Twist system from (5.34) for a time horizon of $T = 8$ by SOS tightening to the Lie-robustified (6.20). The 100 observations


 Figure 6.7: Minimizing x_2 on Flow system (3.9) at order-4 SOS tightening

from this system are pictured in Figure 6.9, yielding a $L = 9$ -dimensional polytope with 34 non-redundant faces. As the order of tightening to program increases from 3 to 4, the red region (level set of $\omega(x)$) tightens to the spiraling attractor region of the $T = 8$ reachable set. The ‘volume’ in the plot titles is not the true volume of the superlevel set $\{x \mid \phi(x) \geq 1\}$, but is instead the Lebesgue estimates $\int_X \phi(x) dx$.

6.7.6 Region of Attraction Example

This example of ROA maximization will concentrate on a controlled version of the Flow dynamics from (3.9) under $L = 6$ inputs

$$\dot{x} = \text{Flow}(x) + \begin{bmatrix} w_1 + w_2 x_1 + w_3 x_2 \\ w_4 + w_5 x_1 + w_6 x_2 \end{bmatrix} \quad (6.47)$$

obeying the polytopic input limits

$$W = \left\{ w \mid \begin{array}{l} \| [w_1; w_4] \|_\infty \leq 0.1, \| [w_2; w_3; w_5; w_6] \|_\infty \leq 0.15 \\ \| [(w_1 + w_2 + w_3); (w_4 + w_5 + w_6)] \|_\infty \leq 0.3 \end{array} \right\}. \quad (6.48)$$

The circle $X_T = \{x \mid 0.1^2 - (x_1 - 0.5)^2 - (x_2 - 0.5)^2 \geq 0\}$ is the destination of the ROA problem with a time horizon of $T = 5$ and a state space of $X = [-1.5, 1.5]^2$. The WSOS tightening of the Lie-robustified problem (6.22) yields bounds for the ROA volume of $d_{2:6}^* = [9.000, 9.000, 6.717, 5.620, 5.187]$.

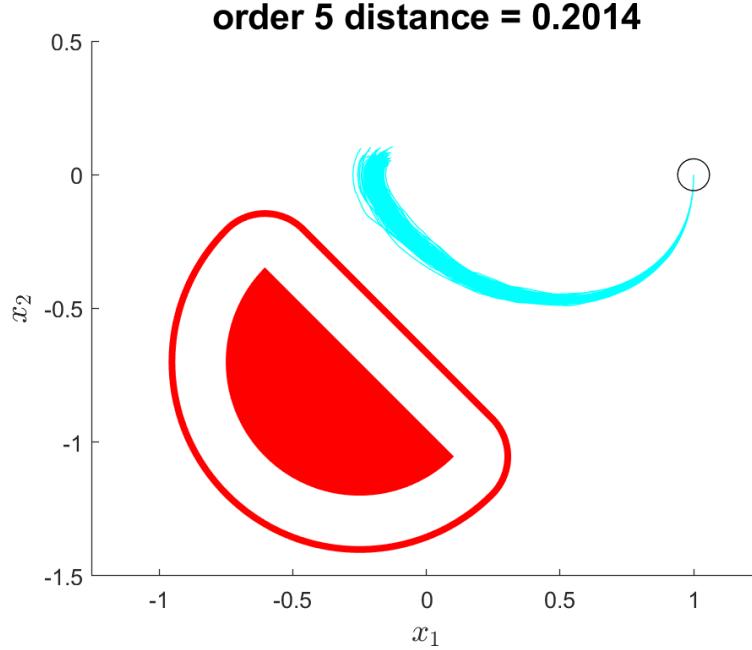


Figure 6.8: Distance estimate of (3.9) at order 5

The destination set X_T is drawn in the black circle in the left subplot of Figure 6.13. The white area is an outer approximation of the true ROA, found as the superlevel set $\{x \mid \phi(x) \geq 1\}$ at order 6. The red area is the sublevel set $\{x \mid \phi(x) \leq 1\}$. The right subplot of Figure 6.13 draws the degree-12 polynomial function $\phi(x)$.

6.8 Conclusion

This work formulated infinite-dimensional robust counterparts decomposing input-affine Lie constraints in analysis and control problems. These robust counterparts may be approximated by continuous multipliers without conservatism under compactness and regularity conditions. Elimination of the noise variables w allows for solution and analysis of formerly intractable problems using the moment-SOS hierarchy. The robust counterpart method was demonstrated on peak estimation, distance estimation, reachable set estimation, and BRS-maximizing control problems. Another environment in which these robust counterparts may be employed is in data-driven systems analysis with affinely-parameterized dictionaries and SDR noise.

The robust counterpart method may be used in other optimization domains with Lie constraints, such as in optimal control input penalties [77] and maximum controlled invariant set

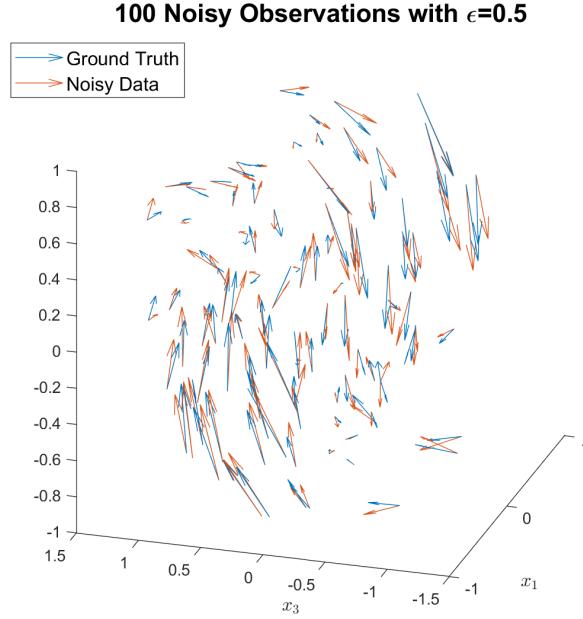


Figure 6.9: 100 observations of Twist system (5.34)

estimation [97].

The Lie robust counterparts depend on input-affine dynamical structure and are therefore restricted to continuous time systems (when nonlinear). One direction is to try and reduce the complexity of analysis of discrete-time dynamics by finding sparse and exploitable structures other than switching. Correlative sparsity is generally incompatible with robust counterparts, but further investigation should lead to cases in which imposing that the multipliers ζ have a CSP [98]. Another avenue is to incorporate warm starts into SDP solvers so system estimates can be updated as more data gets added to \mathcal{D} .

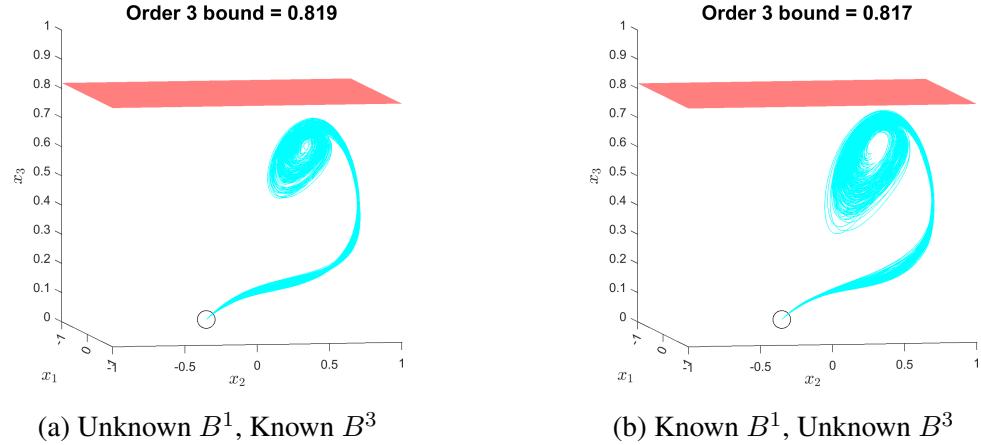


Figure 6.10: Twist (5.34) system where either B^1 or B^3 are unknown.

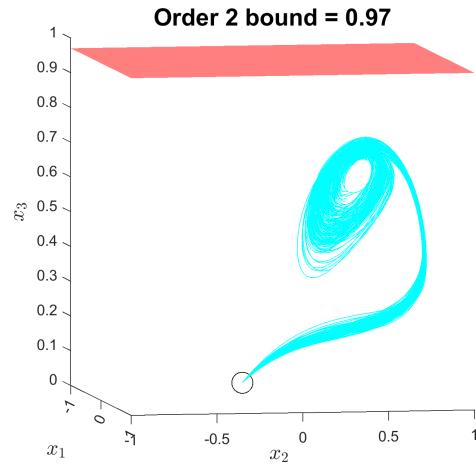


Figure 6.11: Twist (5.34) with unknown (B^1, B^3)

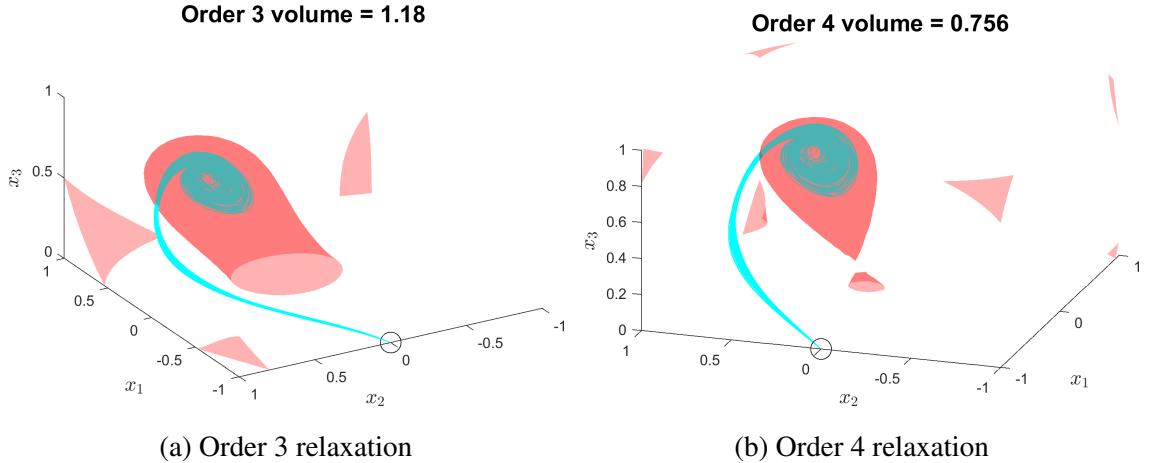


Figure 6.12: Reachable set estimation of twist (5.34) system where B^1 is known and B^3 are unknown.

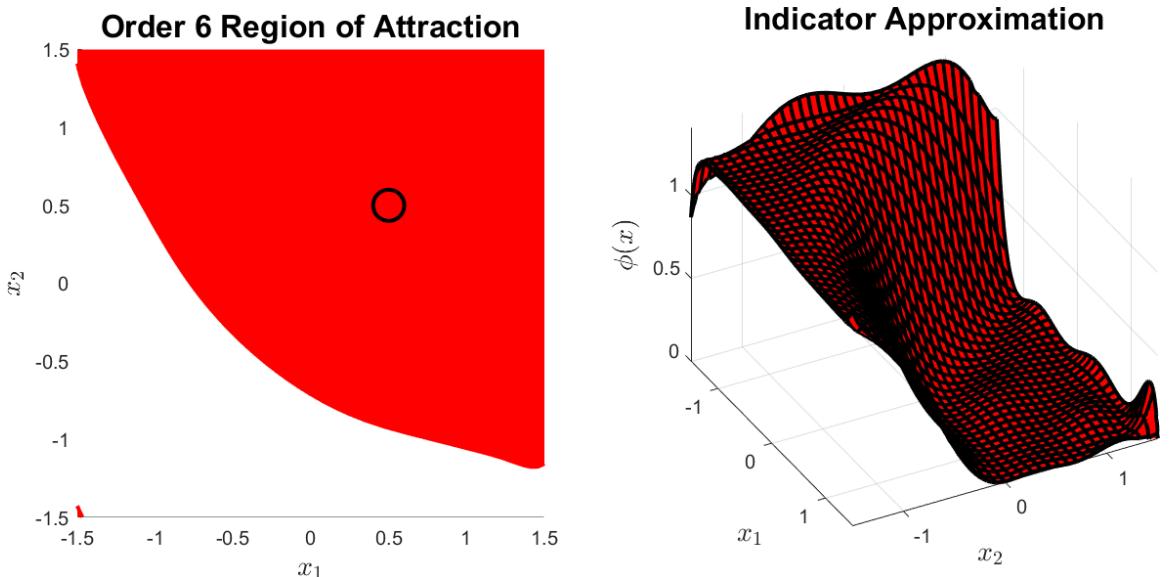


Figure 6.13: Order 6 ROA for controlled Flow (3.9)

Chapter 7

Safety Quantification using Peak Minimizing Control

7.1 Introduction

This chapter adds a third perspective to safety quantification, in addition to the safety margin and distance estimation work from Chapter 5. The safety of trajectories will be quantified by the maximum control effort (OCP cost) needed to crash the agent into the unsafe set. Distance estimation does not tell the full story about safety; a trajectory may lie close to X_u in the sense of distance $c(x; X_u)$, but it could require a severe control effort to render the same trajectory unsafe. An example of this type of safety result is if the tilting of the steering wheel of a car by a maximum extent of 3° over the course of its motion would cause the car to crash. The process of analyzing safety by peak-minimizing-OCP cost will be referred to as ‘crash safety’. This perspective will also be used in the data-driven framework, in which a trajectory is labeled safe if it would require a large constraint violation against any of its state-derivative datapoints in \mathcal{D} in order to crash.

Let $W \subset \mathbb{R}^L$ be a compact input set, and let \mathcal{W} be the class of functions whose graphs satisfy $(t, w(t)) \in [0, T] \times W$. Given a control-cost $J(w)$, we can pose the following peak-

minimizing free-terminal-time OCP:

$$\begin{aligned}
 Q^* &= \inf_{t, x_0, w} \sup_{t' \in [0, t]} J(w(t')) \\
 \dot{x}(t') &= f(t', x(t'), w(t')) \quad \forall t' \in [0, T] \\
 x(t | x_0, w(\cdot)) &\in X_u \\
 w(\cdot) &\in \mathcal{W}, t \in [0, T], x_0 \in X_0.
 \end{aligned} \tag{7.1}$$

The variables of (7.1) are the stopping time t , the initial condition x_0 , and the input process $w(\cdot)$. Assuming for the purposes of this discussion that $J(0) = 0$, $\forall w \neq 0 : J(w) > 0$, and that J possesses connected superlevel sets; the set x_0 is unsafe if $Q^* = 0$ because the process $w(t) = 0$ is sufficient for the trajectory to reach a terminal set of X_u . The value of a nonzero Q^* then measures the amount of control effort (perturbation intensity) needed to render the trajectory unsafe. Connected level sets are imposed to add interpretability to Q^* ; a disconnected choice of J with multiple local minima could yield a large input w with a low Q^* .

A running cost $\int_0^T J(w(t')) dt$ yielding a standard-form (Lagrange) OCP may also be applied, but we elect to use a peak-minimizing cost $\max_{t'} J(w(t'))$ in order to penalize perturbation intensity. The running-cost would penalize a low-magnitude control being applied for an extended period of time, while peak-minimizing control reduces the intensity.

Peak-minimizing control problems, such as in (7.1), are a particular form of robust optimal control in which the minimizing agents are $(t, x_0, w(\cdot))$ and the maximizing agent is $t' \in [0, t]$. Necessary conditions for these robust programs may be found in [99]. Instances of peak-minimizing control include minimizing the maximum number of infected persons in an epidemic under budget constraints [100] and choosing flight parameters to minimize the maximum skin temperature during atmospheric reentry [101, 102]. The work in [12] outlines conversions between peak-minimizing OCPs and equivalent Mayer-form OCPs (terminal cost only).

This chapter transforms program (7.1) into the Mayer OCP using [12], relaxes the nonconvex OCP into an infinite-dimensional LP with the same objective value [7], and then lower-bounds Q^* by using the moment-SOS hierarchy [8, 27]. The robust counterpart method of Chapter 6 will be used to simplify the infinite-dimensional LP when W and the graph of J are both SDR.

This chapter has the following structure: Section 7.2 reviews the peak-minimizing control framework of [12]. Section 7.3 formulates an infinite-dimensional LP to solve (7.1) and calculates a subvalue function to act as a proxy for risk. Section 7.4 applies the crash-safety framework towards L_∞ -penalized data-driven analysis using robust Lie counterparts from Chapter 6. Section 7.5 forms

SOS programs for crash-safety and tabulates their computational complexity. Section 7.6 evaluates the safety of points inside X by a subvalue function of the crash-safety cost. Section 7.7 provides demonstrations of crash-safety. Section 7.8 concludes the chapter. This chapter appears in [103] and was coauthored by Mario Sznaier.

7.2 Peak Minimizing Control

This section reviews the peak-minimizing control problem and a simplified conversion framework based on [12]. Given an objective $\theta : [0, T] \times X \times W \rightarrow \mathbb{R}$, an initial condition x_0 , and a fixed terminal time T , the peak-minimizing control problem is

$$\begin{aligned} P^* = \inf_{w \in \mathcal{W}} \quad & \sup_{t' \in [0, T]} \theta(t', x(t' | x_0, w(\cdot)), w(t')) \\ \dot{x}(t) = f(t', x(t'), w(t')) \quad & \forall t \in [0, T]. \end{aligned} \tag{7.2}$$

The work in [12] details three different methods to convert a peak-minimizing control OCP into a Mayer OCP: pure state constraint, mixed state constraint, and differential inclusion. We will elect to use the first method in [12], which involves the augmentation of constant dynamics by a new state $\dot{z} = 0$:

$$\begin{aligned} P_z^* = \inf_{w \in \mathcal{W}, z \in \mathbb{R}} \quad & z \\ \dot{x}(t) = f(t, x(t), w(t)) \quad & \forall t \in [0, T] \\ \dot{z} = 0 \quad & \forall t \in [0, T] \\ z \geq \theta(t', x(t' | x_0, w(\cdot)), w(t')) \quad & \forall t' \in [0, T]. \end{aligned} \tag{7.3}$$

The parameter z always remains an upper bound on p along trajectories, and the controller $w(\cdot)$ is chosen to reduce this upper bound as much as possible.

Proposition 7.2.1 (Proposition 3.1 of [12]). *The objectives P^* and P_z^* are equal between (7.3) and (7.2).*

7.3 Crash-Safety Program

This section applies peak-minimizing control conversion to the crash-safety task in (7.1).

7.3.1 Motivating Example

This subsection provides an example demonstrating how (7.1) can be used for safety quantification. This example will perturb (3.9) by an uncertainty process restricted to $\forall t : w(t) \in [-1, 1]$:

$$\dot{x} = \begin{bmatrix} x_2 \\ -x_1 - x_2 + \frac{1}{3}x_1^3 \end{bmatrix} + w \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (7.4)$$

Trajectories evolve over a time horizon of $T = 5$ in the state set $X = [-0.6, 1.75] \times [-1.5, 1.5]$ with a maximum corruption of $J_{\max} = 2$. System dynamics are illustrated by the blue streamlines in Figure 7.1. The red half-circle is the unsafe set $X_u = \{x \mid x_2 \leq -0.5, (x_1 - 1)^2 + (x_2 + 0.5)^2 \leq 0.5^2\}$. Two trajectories of this system are highlighted. The green trajectory starts from the top initial point $X_0^1 = [0; 1]$, and the yellow trajectory starts from the bottom initial point $X_0^2 = [1.2966 - 1.5]$. The distance of closest approach to X_u is 0.2498 for both trajectories (matching up to four decimal places). The 0.2498-contour of constant distance is displayed by the red curve surrounding X_u .

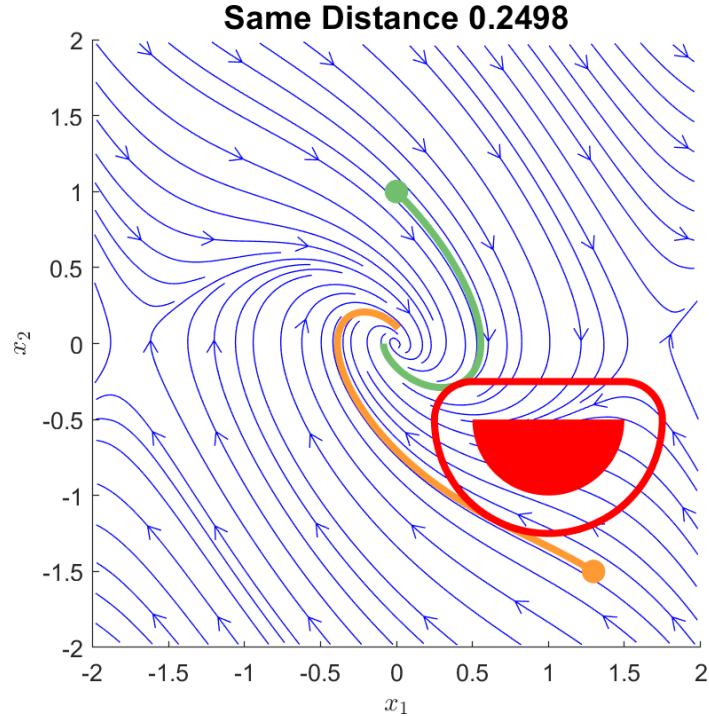


Figure 7.1: Two trajectories with nearly the same distance but different crash-bounds

The OCP solver CasADi [104] returns approximate bounds for (7.1) of $Q^* \approx 0.3160$ for X_0^1 (green) and $Q^* \approx 0.6223$ for X_0^2 (yellow). The points (X_0^1, X_0^2) return nearly identical distances of closest approach, but X_0^2 may be judged as safer than X_0^1 under the disturbance model in (3.9) due to its higher crash-bound value. Degree-4 SOS tightenings of (7.5) developed in the sequel return lower bounds of 0.3018 and 0.5273 respectively.

7.3.2 Assumptions

We will require the following assumptions:

- A1 The sets $[0, T]$, $[0, J_{\max}]$, X, W, X_u, X_0 are all compact.
- A2 The image $f(t, x, W)$ is convex for each fixed (t, x) .
- A3 The dynamics function $f(t, x, w)$ is Lipschitz in the compact domain $[0, T] \times X \times W$.
- A4 If $x(t \mid x_0, w) \in \partial X$ for some $t \in [0, T]$, $x_0 \in X_0$, $w \in \mathcal{W}$, then $x(t' \mid x_0) \notin X \forall t' \in (t, T]$.

7.3.3 Crash-Safety Formulation

We use the peak-minimizing control conversion of [12] on program (7.1):

Theorem 7.3.1. *The following program has the same optimal value as (7.1)*

$$Q_z^* = \inf_{t, x_0, z, w} z \quad (7.5a)$$

$$\dot{x}(t') = f(t', x(t'), w(t'), \dot{z}(t')) = 0 \quad \forall t' \in [0, T] \quad (7.5b)$$

$$J(w(t')) \leq z \quad \forall t' \in [0, T] \quad (7.5c)$$

$$x(t \mid x_0, w(\cdot)) \in X_u \quad (7.5d)$$

$$w(\cdot) \in W, t \in [0, T], x_0 \in X_0, z \in [0, J_{\max}]. \quad (7.5e)$$

Proof. This follows from Proposition 7.2.1 under the following changes:

1. Free terminal time $t \in [0, T]$
2. Terminal state constraint in X_u
3. Input-state constraint in (7.5c)

4. Free initial condition $x_0 \in X_0$

The state z upper-bounds the worst-case control $\sup_{t' \in [0, t]} J(w(t'))$, yielding the peak-control-minimized cost $Q_z^* = Q^*$. The work in [12] allows the peak-minimized term to be minimized to contain u . \square

7.3.4 Linear Programs

Define the following compact support sets involving w and z

$$Z = [0, J_{\max}] \quad \Omega = \{(w, z) \in W \times Z : J(w) \leq z\}. \quad (7.6)$$

Let \mathcal{L}_f be the Lie derivative associated with f for $v(t, x, z) \in C^1$ as

$$\mathcal{L}_f v(t, x, z, w) = (\partial_t + f(t, x, w) \cdot \nabla_x) v(t, x, z). \quad (7.7)$$

A measure LP for (7.5) in terms of an initial measure μ_0 , terminal measure μ_u , and occupation measure μ is

$$q^* = \inf_{\mu_0, \mu_p, \mu} \langle z, \mu_u \rangle \quad (7.8a)$$

$$\mu_u = \delta_0 \otimes \mu_0 + \pi_\#^{txz} \mathcal{L}_f^\dagger \mu \quad (7.8b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (7.8c)$$

$$\mu_0 \in \mathcal{M}_+(X \times Z) \quad (7.8d)$$

$$\mu_u \in \mathcal{M}_+([0, T] \times X_u \times Z) \quad (7.8e)$$

$$\mu \in \mathcal{M}_+([0, T] \times X \times \Omega). \quad (7.8f)$$

Constraint (7.8a) is a Liouville equation in the sense of (3.6).

Lemma 7.3.2. *There exists a feasible solution to (7.8b)-(7.8f) under A1-A4*

Proof. Let $t^* \in [0, T]$ be a stopping time, $x_0 \in X_0$ be an initial condition, and $w(\cdot) \in \mathcal{W}$ be an input such that $x(t^* | x_0, w(\cdot)) \in X_u$. Let z^* be a feasible solution to $\forall t \in [0, t^*] : (z^*, w(t)) \in \Omega$. Then the probability measures can be set to $\mu_0 = \delta_{x=x_0, z=z^*}$ and $\mu_u = \delta_{t=t^*, x=x(t^* | x_0, w(\cdot)), z=z^*}$, and μ can be assigned to the occupation measure of $t \mapsto (t, x(t^* | x_0, w(\cdot)), w(t))$ in the times $[0, t^*]$. \square

Remark 7.3.1. *The process of 7.3.2 to generate a feasible measure solution may be used when only A1 and A4 are active, thus certifying that $m^* \leq Q^*$.*

Theorem 7.3.3. *Under assumptions A1-A5, programs (7.1) and (7.8) will have equal objectives $q^* = Q^*$.*

Proof. Program (7.5) with optimum Q_z^* is a standard-form OCP with free terminal time and zero running cost. Under assumptions A1-A5, Theorem 2.1 of [7] proves that $Q_z^* = q^*$. Section 6.3 of [7] specifically discusses state-dependent controls (e.g. $(w, z) \in \Omega$). Theorem 7.3.1 provides that $Q^* = Q_z^*$, which together implies that $Q^* = q^*$. \square

An auxiliary function $v \in C^1$ may be defined to form a dual LP to (7.8) as

$$d^* = \sup_{\gamma \in \mathbb{R}, v} \gamma \quad (7.9a)$$

$$v(0, x, z) \geq \gamma \quad \forall (x, z) \in X_0 \times Z \quad (7.9b)$$

$$v(t, x, z) \leq z \quad \forall (t, x, z) \in [0, T] \times X_u \times Z \quad (7.9c)$$

$$\mathcal{L}_f v(t, x, z, w) \geq 0 \quad \forall (t, x, z, w) \in [0, T] \times X \times \Omega \quad (7.9d)$$

$$v(t, x, z) \in C^1([0, T] \times X \times Z). \quad (7.9e)$$

Theorem 7.3.4. *Strong duality occurs with $q^* = d^*$ between (7.8) and (7.9) under assumptions A1-A5.*

Proof. This holds by standard OCP LP duality arguments from [105, 7, 16]. \square

7.4 Robust Crash-Safety and Data-Driven Analysis

This section motivates crash-safety in the context of data-driven analysis. This section will remove the restriction that the performance function satisfies $J(0) = 0$, but will retain the property that the level sets of J are connected.

7.4.1 Data-Driven Overview

We recall the data-driven setting of Section 6.6. We are given data $\mathcal{D} = \{(t_k, x_k, y_k)\}_{k=1}^{N_s}$ and a dictionary of functions $(f_0, \{f_\ell\}_{\ell=1}^L)$ according to dynamics (6.1). There exists at least one ground-truth choice of parameters $w^* \in \mathbb{R}^L$ such that

$$F(t, x) = f_0(t, x) + \sum_{\ell=1}^L w_\ell^* f_\ell(t, x). \quad (7.10)$$

In the L_∞ -bounded polytopic framework, the crash-safety problem (7.5) aims to find a minimum upper bound on data corruption in any data record needed to crash into the unsafe set:

$$Z^* = \inf_{t, x_0, z, w} z \quad (7.11a)$$

$$\dot{x}(t') = f_0(t', x) + \sum_{\ell=1}^L w_\ell f_\ell(t', x(t')), \quad \dot{z}(t') = 0 \quad \forall t' \in [0, T] \quad (7.11b)$$

$$x(t \mid x_0, w) \in X_u \quad (7.11c)$$

$$x(0) \in X_0 \quad (7.11d)$$

$$z \geq \|f_0(t_k, x_k) + \sum_{\ell=1}^L w_\ell f_\ell(t_k, x_k) - y_k\|_\infty \quad \forall k = 1..N_s \quad (7.11e)$$

$$z \in [0, J_{\max}], \quad w \in \mathbb{R}^L, \quad t \in [0, T]. \quad (7.11f)$$

If the returned value of (7.11) is $Z^* = 0$, then there exists some choice of model parameters w that exactly fit the data \mathcal{D} such that at least one trajectory $x(\cdot)$ starting from X_0 is unsafe (crashes into X_0). Values of Z^* greater than 0 are a certificate of safety in the model structure. A larger value of Z^* indicates that the data must be increasingly corrupted in order to render any trajectory unsafe.

7.4.2 Robust Data-Driven Program

The state constraint (7.11e) defines a z -scaled polytope W from (6.44). The data-derived constant matrices $\Gamma \in \mathbb{R}^{2nT \times L}$, $h \in \mathbb{R}^{2nT}$ from (6.43) may be used describe constraint (7.11e) as

$$\Gamma w(t) \leq z\mathbf{1} + h \quad (7.12)$$

The performance function of the L_∞ -data-driven analysis is

$$J(w) = \max_j (h - \Gamma w)_j. \quad (7.13)$$

The true corruption value is $J(w^*)$.

Given a maximum upper-bound on data corruption J_{\max} , the support sets (Z, Ω) may be defined for the data-driven case as

$$Z = [0, J_{\max}] \quad \Omega = \{(w, z) \in \mathbb{R}^L \times Z : \Gamma w \leq z\mathbf{1} + h\}. \quad (7.14)$$

The support set (7.14) is polytopic in the uncertainty w .

Theorem 7.4.1. *The Lie constraint in (7.9d) may be robustified in a nonconservative manner into*

$$\mathcal{L}_{f_0} v - (z\mathbf{1} + h)^T \zeta \geq 0 \quad \forall (t, x, z) \in [0, T] \times X \times [0, J_{\max}] \quad (7.15a)$$

$$(\Gamma^T)_\ell \zeta + f_\ell \cdot \nabla_x v = 0 \quad \forall \ell = 1..L \quad (7.15b)$$

$$\zeta_j \in C_+([0, T] \times X \times Z) \quad \forall j = 1..2nT. \quad (7.15c)$$

Proof. See Theorem 6.4.2 for the robustification. We perform the following assignments to the robust inequality in (6.6):

$$a_0 = \emptyset \quad a_\ell = \emptyset \quad \forall \ell = 1..L \quad (7.16a)$$

$$b_0 = \mathcal{L}_{f_0} v \quad b_\ell = f_\ell \cdot \nabla_x v \quad \forall \ell = 1..L \quad (7.16b)$$

$$A = -\Gamma \quad G = \emptyset \quad (7.16c)$$

$$e = z + h \quad K = \prod_{s=1}^{2nT} \mathbb{R}_{\geq 0}. \quad (7.16d)$$

The parameters of this problem are $(t, x, z) \in [0, T] \times X \times Z$. The elements e, b_0, b_ℓ are all continuous functions of the parameters, and the constraint matrix A is constant in the parameters. \square

7.5 SOS Programs

This section poses finite-dimensional SOS tightenings to the infinite-dimensional crash-safety programs.

We will require a strengthening of assumption A1:

A5 The sets $(X, X_u, X_0, [0, T], Z, \Omega)$ are all Archimedean BSA sets and the dynamics $f(t, x, w)$ are polynomial.

7.5.1 Standard Crash-Safety

For a given degree d , define $\tilde{d} = d + \lfloor \deg f / 2 \rfloor$ as the dynamics degree of $f(t, x, w)$. The degree- d SOS tightening of program (7.9) is

$$q_d^* = \max_{\gamma \in \mathbb{R}, v} \gamma \quad (7.17a)$$

$$v(0, x, z) - \gamma \in \Sigma[X_0 \times Z]_{\leq d} \quad (7.17b)$$

$$z - v(t, x, z) \in \Sigma[[0, T] \times X_u \times Z]_{\leq d} \quad (7.17c)$$

$$\mathcal{L}_f v(t, x, z, w) \in \Sigma_{\tilde{d}}[[0, T] \times X \times \Omega] \quad (7.17d)$$

$$v(t, x, z) \in \mathbb{R}[t, x, z]_{\leq 2d}. \quad (7.17e)$$

We will show that all measures in (7.8) are bounded in order to prove convergence of (7.17) to the true value in (7.11).

Lemma 7.5.1. *All measures (μ_0, μ_u, μ) in (7.8) are bounded under A1-A5.*

Proof. All support sets are compact by assumption A1. The measure μ_0 has mass 1 by (7.8c). Substitution of $v(t, x, z) = 1$ into (7.8b) results in $\langle 1, \mu_u \rangle = \langle 1, \mu_0 \rangle = 1$, and applying $v(t, x, z) = t$ yields $\langle 1, \mu \rangle = \langle t, \mu_u \rangle \leq T$. \square

Theorem 7.5.2. *Under assumptions A1-A5, then the sequence of bounds q_d^* will converge as $\lim_{d \rightarrow \infty} q_d^* = Q^*$ to the optimum of (7.1).*

Proof. This convergence will occur by Corollary 8 of [23], along with convergence in Theorem 7.3.3, boundedness of measures in 7.5.1, the infinite-dimensional strong duality Theorem 7.3.4, and strong duality between their finite-dimensional SDP truncations [21, Arguments from Theorem 4]. \square

7.5.2 Robust Crash-Safety

We now apply the robust counterpart from (7.15) to (7.9) in order to form an SOS program for the L_∞ data-driven scenario. Define $\tilde{d} = d + \max_{\ell \in 0..L} \lfloor \deg f_\ell / 2 \rfloor$ as the dynamics degree of (6.1). The L_∞ -bounded data-driven robust crash-safety SOS tightening at degree d is

$$\tilde{q}_d^* = \max_{\gamma \in \mathbb{R}, v} \gamma \quad (7.18a)$$

$$v(0, x, z) - \gamma \in \Sigma_d[X_0 \times Z] \quad (7.18b)$$

$$z - v(t, x, z) \in \Sigma_d[[0, T] \times X_u \times Z] \quad (7.18c)$$

$$\mathcal{L}_{f_0} v - (z \mathbf{1} + h)^T \zeta \in \Sigma_{\tilde{d}}[[0, T] \times X \times Z] \quad (7.18d)$$

$$\text{coeff}_{txz}(-(\Gamma^+)_\ell \zeta + f_\ell \cdot \nabla_x v) = 0 \quad \forall \ell = 1..L \quad (7.18e)$$

$$v(t, x, z) \in \mathbb{R}[t, x, z]_{\leq 2d} \quad (7.18f)$$

$$\zeta_j \in \Sigma[[0, T] \times X \times Z]_{\leq \tilde{d}-1} \quad \forall j = 1..2nT. \quad (7.18g)$$

Theorem 7.5.3. *Under assumptions A1-A5 and assuming L_∞ noise structure, the sequence of optimal values from (7.18) will converge as $\lim_{d \rightarrow \infty} \tilde{q}_d^* = Q^*$.*

Proof. The Lie constraint may be robustified by Theorem (7.4.1). The SOS program in (7.18) will converge to a strict version of (7.9) by Theorem (6.5.4) under the polynomial v restriction. Strictness is not overly restrictive when performing smooth approximations, as shown in the proof of Proposition 5 in [106]. \square

Remark 7.5.1. *The degree of ζ in (7.18g) is set to $2(\tilde{d}-1)$ so as to ensure that $\deg z \mathbf{1}^T \zeta = 2\tilde{d}-1 \leq 2\tilde{d}$ in (7.18d).*

7.5.3 Computational Complexity

Program (7.17) has three WSOS constraints, leading to Gram matrices of maximal size $\binom{n+1+d}{d}$, $\binom{n+2+d}{d}$, $\binom{n+L+2+\tilde{d}}{\tilde{d}}$. The performance of SDPs derived from (7.17) is dominated by the largest size $\binom{n+L+2+\tilde{d}}{\tilde{d}}$ and scales as $(n + L + 2)^{6\tilde{d}}$ or $\tilde{d}^{4(n+L+2)}$ (per-iteration complexity of interior point methods moment-SOS).

The robustified program in (7.18) breaks up the Lie constraint's maximal-size Gram matrix dimension $\binom{n+L+2+\tilde{d}}{\tilde{d}}$ into one matrix of size $\binom{n+2+\tilde{d}}{\tilde{d}}$ (7.18d) and $2nT$ Gram matrices of size $\binom{n+1+\tilde{d}}{\tilde{d}}$ (7.18g).

The nonredundant face identification method of [94] requires caution when attempting to reduce complexity of (7.18). Faces of W that are active at $z = z_1$ may no longer be active at $z = z_2 \geq z_1$ or vice versa. A bound on (7.18) computed using a subset of faces (constraints) in \mathcal{D} will necessarily be lower than using all faces. This conservatism can be reduced while still eliminating faces by taking the union of active faces of the polytopes in w from (7.12) at a set of values $z \in [0, J_{\max}]$.

7.6 Subvalue Map

Program (7.9) returns the worst-case crash safety over a set of initial conditions X_0 . We briefly discuss an extension of the crash-safety technique to assessing the safety of arbitrary initial conditions.

7.6.1 Value Functions

We define the fixed- z value function of (7.5) (when starting at $X_0 = x'$) as

$$V(x', z) = \begin{cases} z & z \in [0, J_{\max}], \exists t \in [0, T], w(\cdot) \in \mathcal{W} \mid \forall t' \in [0, t] \\ & x(t \mid x_0, w(\cdot)) \in X_u \\ \infty & \text{otherwise.} \end{cases} \quad (7.19)$$

The value function $V(x', z)$ is infinite if the control problem of steering a point from x' to X_u is infeasible within the performance budget $J(w) \leq z$. The value function of (7.5) when

restricted to the single initial condition x' is

$$Q(x') = \inf_{z \in [0, J_{\max}]} V(x', z). \quad (7.20)$$

The value function $Q(x')$ will have an upper bound of J_{\max} if $Q(x')$ is finite, and otherwise will have a value of ∞ . We make no assumptions of continuity or boundedness of $Q(x')$, beyond A1's assurance that J_{\max} is finite.

7.6.2 Subvalue Approximations

We now use the moment-SOS hierarchy to develop subvalue maps to lower-bound $Q(x')$ from (7.20).

Proposition 7.6.1. *Any function $v(t, x, z)$ that satisfies (7.9c) and (7.9d) obeys $v(0, x, z) \leq V(x', z)$ from (7.19) at all $(x, z) \in X \times Z$.*

Proof. Equations (7.9c) and (7.9d) are inequality constrained versions of the Hamilton-Jacobi-Bellman equality constraints for an optimal value function v^* [107]:

$$v^*(t, x', z) = z \quad \forall (t, x', z) \in [0, T] \times X_u \times Z \quad (7.21a)$$

$$\min_{w|(w, z) \in \Omega} \mathcal{L}_f v^*(t, x', z, w) = 0 \quad \forall (t, x', z) \in [0, T] \times X \times Z. \quad (7.21b)$$

Refer to the Section 4 of [8] and the proof of Proposition 1 of [106] for the establishment of subvalue relations. \square

Let $\varphi \in \mathcal{M}_+(X)$ be a probability distribution with easily computable moments (e.g., uniform distribution over X when X is a ball or a box), and $Q_{\max} \geq J_{\max}$ be a finite control cap.

Theorem 7.6.2. *The following program provides a subvalue function $q(x) \leq Q(x)$:*

$$J^* = \sup \int_X q(x) d\varphi(x) \quad (7.22a)$$

$$q(x) \leq v(0, x, z) \quad \forall (x, z) \in X \times [0, Z_{\max}] \quad (7.22b)$$

$$q(x) \leq Q_{\max} \quad \forall x \in \text{supp}(\varphi) \quad (7.22c)$$

$$z \geq v(t, x, z) \quad \forall (t, x, z) \in [0, T] \times X_u \times Z \quad (7.22d)$$

$$\mathcal{L}_f v(t, x, z, w) \geq 0 \quad \forall (t, x, z, w) \in [0, T] \times X \times \Omega \quad (7.22e)$$

$$v \in C^1([0, T] \times X \times Z) \quad (7.22f)$$

$$q \in C(X). \quad (7.22g)$$

Proof. Proposition 7.6.1 proves that $v(0, x, z) \leq V(x, z)$ from (7.19). Constraint (7.22b) imposes that $q(x) \leq v(0, x, z) \leq V(x, z)$ for all $x \in X$, which implies that $q(x) \leq \inf_z v(0, x, z)$ for all $x \in X$. From the definition of $Q(x')$ in (7.20) with $Q(x') \leq \inf_z V(x', z)$, it therefore holds that $q(x) \leq Q(x)$ for all $x \in X$. \square

Corollary 1. *The objective J^* from (7.22) is finite and is bounded above by $J^* \leq Q_{\max}$.*

Proof. Constraint (7.22c) requires that $q(x)$ is upper-bounded by Q_{\max} . The objective (7.22a) is therefore upper-bounded by

$$\int_X q(x)d\varphi(x) \leq \int_X Q_{\max}d\varphi(x) \leq Q_{\max} \int_X d\varphi(x) = Q_{\max}, \quad (7.23)$$

given that φ is a probability distribution. \square

Remark 7.6.1. *Let v be a subvalue solution to (7.22d)-(7.22f). Any point $x' \in X$ such that $\inf_{z \in Z} v(0, x', z) > J_{\max}$ implies that $Q(x') = \infty$.*

Remark 7.6.2. *Without the Q_{\max} cap in (7.22c), the optimal value of (7.22) could be $J^* = \infty$ if $\exists x' \in X : Q(x') = \infty$.*

Define $q_d \in \mathbb{R}[x]_{\leq 2d}, v_d \in \mathbb{R}[t, x, z]_{\leq 2d}$ as the polynomials obtained by solving the degree- d SOS tightening of (7.22). Let $I_u(x)$ be the indicator function

$$I_u(x) \begin{cases} 0 & x \in X_u \\ -\infty & x \notin X_u \end{cases}. \quad (7.24)$$

For a sequence of orders $d' = 1..d$, a parametric function $q_{1:d}$ may be defined as

$$q_{1:d}(x) = \max(I_u(x), \max_{d' \in 1..d} q_{d'}(x)). \quad (7.25)$$

Definition 7.6.1 ([60]). *A sequence of continuous functions $\{q_k(x)\}$ converges **almost uniformly** to $Q(x)$ with respect to a measure $\varphi \in \mathcal{M}_+(X)$ if $\epsilon > 0 : \exists A \subseteq X$, such that $q_k \rightarrow Q$ uniformly on $X \setminus A$ and $\varphi(A) < \epsilon$.*

Theorem 7.6.3. *The function $q_{1:d}(x)$ will converge almost uniformly to $\min(Q_{\max}, Q(x))$ on the state-space $\text{supp}(\varphi) \in X$.*

Proof. Let $\tilde{v} \in \mathbb{R}[t, x, z]$ be a polynomial subvalue function that obeys (7.22d)-(7.22f). Corollary 2.5 of [60] proves that the parameterized program $q_{1:d}$ will converge φ -almost uniformly to $\min(Q_{\max}, \min_z \tilde{v}(0, x, z))$, resulting in

$$\lim_{k \rightarrow \infty} \int_X |Q(x) - q_k(x)| d\varphi(x) = 0. \quad (7.26)$$

Increasing the degree of sublevel polynomials \tilde{v} allows for the choice of admissible \tilde{v} such that $\tilde{v}(0, x, z)$ converges in an L^1 -sense to $V(x, z)$ whenever $V(x, z) \leq Q_{\max}$ [106, Propositions 5 and 6], thus proving the theorem. \square

Remark 7.6.3. *The subvalue approximation $q_{1:d}$ in (7.25) is vulnerable to a Gibbs phenomenon, which is common among all polynomial optimization methods [108, 109]. Remark 7.6.1 is vital in establishing infeasibility of reaching X_u , but choosing $Q_{\max} = J_{\max}$ may lead to Gibbs phenomena that distort the infeasibility $Q(x') = \infty$ into $q(x') \leq J_{\max}$. Picking $Q_{\max} > J_{\max}$ (such as $Q_{\max} = 4J_{\max}$) allows for slack in the range of $[Q_{\max} - J_{\max}, Q_{\max}]$, which hopefully could contain the Gibbs phenomena when establishing infeasibility (safety up to J_{\max}).*

Remark 7.6.4. *The Lie constraint in (7.22e) may be robustified through the methods in 7.4 to produce data-driven subvalue maps.*

7.7 Examples

This section demonstrates the utility of the crash-safety framework. Robust decompositions of the Lie constraint are applied in all examples. MATLAB R2021a code to generate examples is available at <https://github.com/Jarmill/crash-safety>. All SDP are generated using YALMIP [68] and solved using Mosek [49]. Finite-degree crash-bounds from (7.18) are compared against OCP bounds found using the solver CasADi [104].

7.7.1 Single-Input Subvalue Comparison

This example demonstrates the computation of crash-bounds and the creation of crash-subvalue functionals for system (7.4) with $J_{\max} = 1$ and $Q_{\max} = 4$. This subvalue is constructed by solving SOS tightenings of (7.22) in the space $X = [-2, 2]^2$ and in the time horizon $t \in [0, 5]$

7.7.1.1 Half-Circle

The first part of this example involves the half-circle respect to the unsafe set $X_u = \{x \mid (x_1 + 0.25)^2 + (x_2 + 0.7)^2 \leq 0.5^2, (0.95 + x_1 + x_2)/\sqrt{2} \leq 0\}$. Figure 7.2 draws the unsafe set X_u in red. The color shading (colorbar) plots $q_{1:5}(x)$ clamped to the range $[0, J_{\max}] = [0, 1]$. The integral objective values of SOS tightening (7.22) at degrees 1..5 are $J_{1:5}^* = [1.934 \times 10^{-7}, 4.864 \times 10^{-7}, 3.0794, 5.992, 8.260]$.

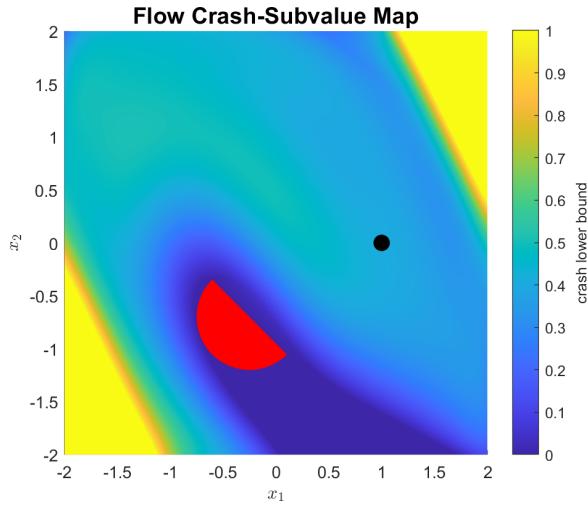


Figure 7.2: Subvalue function for Flow system (7.4) between degrees 1..5.

The black dot in Figure 7.2 is the specific initial point $X_0 = [1; 0]$. Table 7.1 lists crash-bounds on (7.4) starting at X_0 . The subvalue bound (7.22) is lower than the corresponding degree bounds at the X_0 -specific program (7.9).

Table 7.1: Crash-bounds at $X_0 = [1; 0]$ under SOS tightenings

order	1	2	3	4	5
subvalue (7.22)	1.089×10^{-9}	1.607×10^{-9}	0.1473	0.3392	0.4053
specific (7.9)	1.117×10^{-7}	0.1843	0.4369	0.5092	0.5118

We now consider worst-case crash-bounds for the half-circle set with respect to the perturbed flow system (7.4) and the circular initial set $X_0 = \{x \mid 0.4^2 \geq (x_1 - 1)^2 + x_2\}$. Crash-bounds as computed by (7.18) (SOS tightenings to (7.9)) in degrees 1..5 are $[8.101 \times 10^{-8}, 6.590 \times 10^{-2}, 0.4054, 0.4631, 0.4638]$. The degree-5 lower-bound of 0.4638 should be compared against the numerical bound of 0.4639 produced by CasADi. The numerically solved trajectory (blue curve)

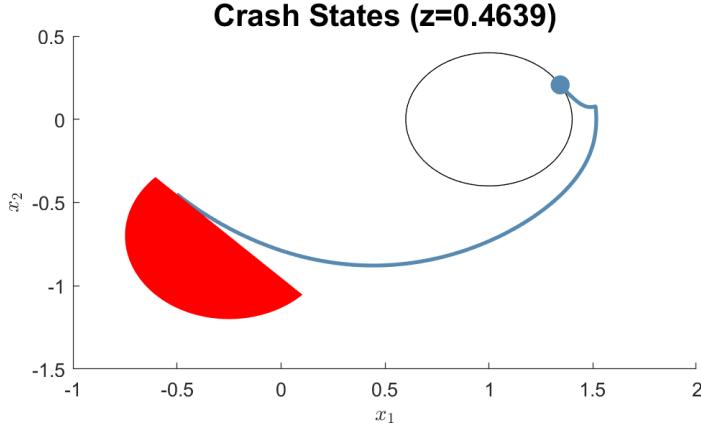


Figure 7.3: Numerical optimal control yields worst-case $Q^* \approx 0.4639$ for the half-circle X_u

is plotted in Figure 7.3, along with the unsafe set X_u (red half-circle) and the initial set X_0 (black circle). The initial point of the controlled trajectory (blue dot) is $x_0 \approx [1.3424; 0.2069]$.

7.7.1.2 Moon

The second part of this example has a nonconvex moon-shaped unsafe set

$$X_u = \{x \mid 0.8^2 - (x_1 - 0.4)^2 - (x_2 + 0.4)^2 \geq 0, (x_1 - 0.6596)^2 + (x_2 - 0.3989)^2 - 1.16^2 \geq 0\}. \quad (7.27)$$

Figure 7.4 displays a controlled trajectory (blue curve) starting from $X_0 = [0; 0]$ (black circle) and terminating in the X_u (red moon), as computed by CasADI.

Table 7.2 lists subvalue (7.22) and specific (7.9) crash-bounds for $X_0 = [0; 0]$ between degrees 1..5. The objectives of the SOS tightenings to (7.22) are $J_{1..5}^* = [1.973 \times 10^{-7}, 1.323 \times 10^{-7}, 1.027, 3.188, 4.502]$.

Table 7.2: Crash-bounds at $X_0 = [0; 0]$ for the moon (7.27) under SOS tightenings

order	1	2	3	4	5
subvalue (7.22)	8.770×10^{-9}	4.652×10^{-10}	-7.861×10^{-2}	-5.692×10^{-3}	7.721×10^{-2}
specific (7.9)	2.723×10^{-8}	0.1010	0.2912	0.3216	0.3224

Figure 7.5 plots the subvalue function $q_{1:5}(x)$ from (7.25) under a cap of $Q_{\max} = 2$ (and $J_{\max} = 1$). All values of $q_{1:5}$ in Figure 7.5 are clamped to $[0, J_{\max}]$.

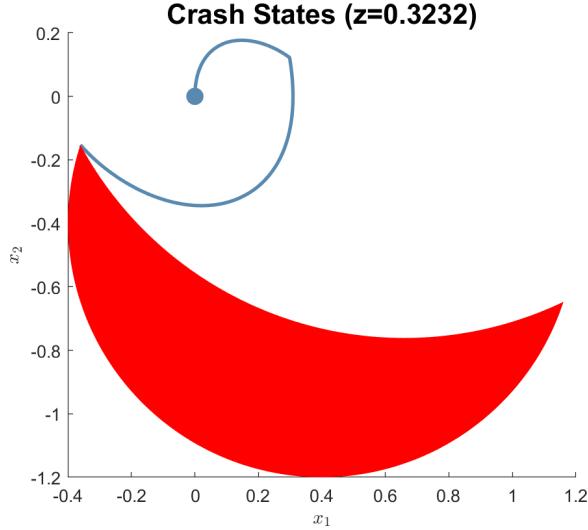


Figure 7.4: Numerical optimal control yields $Q^* \approx 0.3232$ for the moon X_u

7.7.2 Data-Driven Flow System

Data is collected for the Flow system (3.9) from $N = 40$ samples with a ground-truth noise bound of $\epsilon = 0.5$ in the coordinate \dot{x}_2 . The parameterized polytope $\Omega_z = \{w \mid Aw \leq b + z\}$ (Ω with fixed z value) has $L = 10$ dimensions and $m = 2nT = 80$. The minimum possible corruption while obeying (6.37) under the cubic noise model is $\inf_{(w,z) \in \Omega} z = 0.4617$.

The crash-safety problem (7.9) and subvalue problem (7.22) were solved with the unsafe set $X_u = \{x \mid (x_1 + 0.25)^2 + (x_2 + 0.7)^2 \leq 0.5^2, (0.95 + x_1 + x_2)/\sqrt{2} \leq 0\}$ between $t = [0, 5]$ time units in the space $X = \{x \in \mathbb{R}^2 : \|x\|_2^2 \leq 8\}$. The subvalue problem (7.22) integrates over the uniform measure of the ball X .

Table 7.3 reports bounds for the crash-corruption $Q(X_0)$ by solving Lie-robustified SOS tightenings of (7.9) and (7.22) from degrees 1..4 with $J_{\max} = 1$, $Q_{\max} = 4$. The objective function (integrals of $q(x)$) for the subvalue (7.22) are $J_{1:4}^* = [0.2193, 3.8185, 7.8326, 18.5945]$. The subvalue-estimated control cost at X_0 between degrees 1..4 is 0.3399 by Equation (7.25). The subvalue-estimated bound is valid for all $x \in X$, and is therefore lower than the bound $q_4^* = 0.5499$ from (7.9) that focuses exclusively on the initial point X_0 .

Figure 7.6 plots the subvalue function from (7.25) on the data-driven flow system. Subvalues in the plot are clamped to the range $[0, J_{\max}] = [0, 1]$.

Safety of trajectories starting in X_0 is certified because the crash-bound $\tilde{q}_4^* = 0.5499$ is greater than the ground-truth noise-bound $\epsilon = 0.5$. Figure 7.7 uses the CasADi optimal control suite

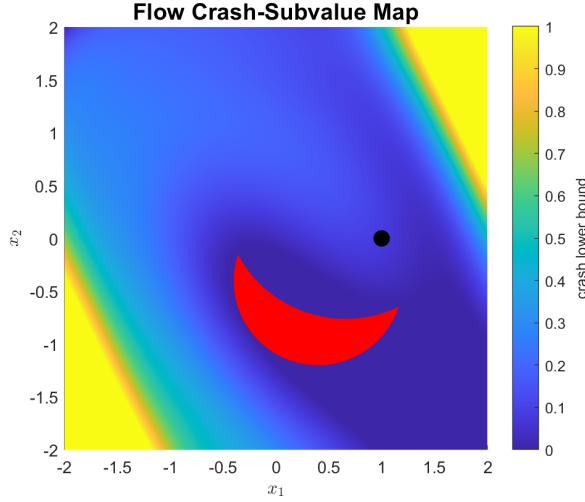


Figure 7.5: Subvalue map for the moon (7.27) on the flow system (7.4)

Table 7.3: Data-Driven Crash-bounds at $X_0 = [1; 0]$ under SOS tightenings

order	1	2	3	4
specific (7.9)	0.0582	0.4423	0.4864	0.5499
subvalue (7.22)	6.180×10^{-3}	0.1829	0.3399	-19.01

[104] to numerically solve the crash program (7.1). The numerical crash-bound of $q^{\text{CasADI}} = 0.5499$ is approximately equal (up to four decimal places) to the crash-bound $q_4^* = 0.5499$.

The left plot of figure 7.8 shows the applied control of the $L = 10$ inputs. The right plot demonstrates how the polytopic input constraint is obeyed with respect to the crash bound $q^{\text{CasADI}} = 0.5499$ (upper and lower black lines).

These crash-bounds should be compared against the L_2 distance estimates of $c_{1:5}^* = [1.698 \times 10^{-5}, 0.1936, 0.2003, 0.2009, 0.2013]$ from Section 6.7.3. The distance estimates do not indicate that adding an additional budget of 0.0499 constraint violation will cause at least trajectory to enter the unsafe set.

7.8 Conclusion

This chapter utilized peak minimizing control in order to perform safety analysis. Crash-safety adds a new perspective on the safety of trajectories, covering some of the blind spots of distance estimation and safety margins. Crash-safety may be applied in the context of data-driven systems analysis by quantifying the minimum tolerable corruption in a noise model before a trajectory

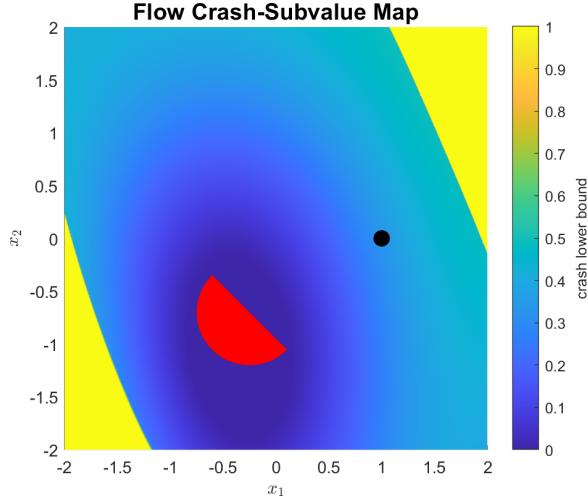


Figure 7.6: Subvalue for data-driven (3.9) between degrees 1..4

is at risk of being unsafe.

Future work involves attempting to reduce computational complexity of the Crash programs (7.17) by identifying new kinds of structure (in addition to SDR robust decompositions) to hopefully allow for real-time computation. Other extensions could include applying these methods to other classes of systems (e.g., discrete-time, hybrid), and creating a stochastic interpretation of crash-safety.

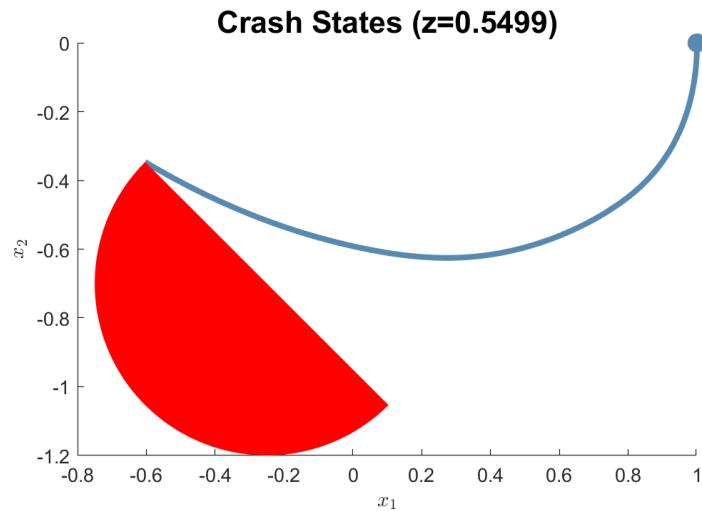


Figure 7.7: Numerically computed crash-bound for data-driven Flow (3.9)

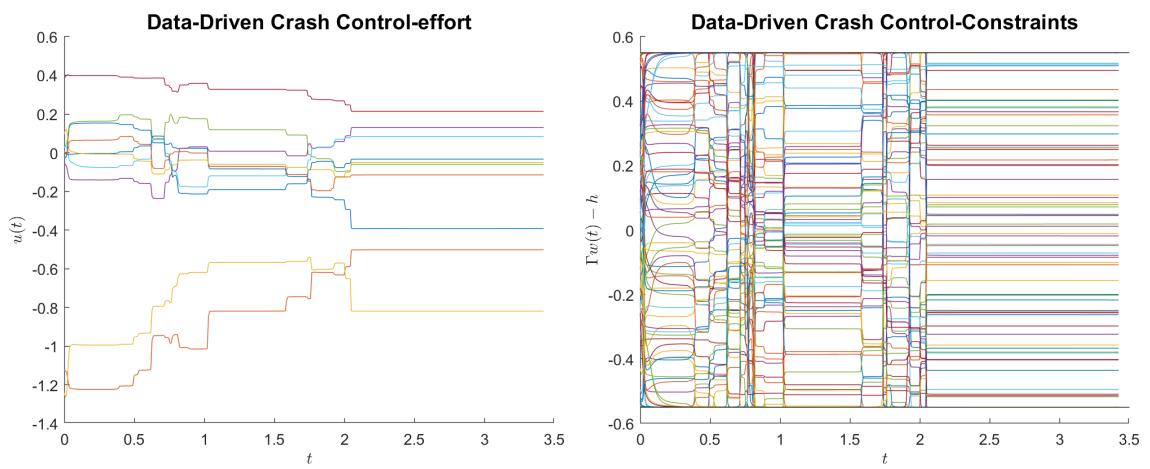


Figure 7.8: Applied control for the data-driven Flow crash system.

Chapter 8

Distance-Maximizing Control

This chapter addresses the control problem of steering from an initial set X_0 to a terminal set X_T in a specified time $t \in [0, T]$ while maximizing the distance of closest approach to an unsafe set X_u . Chapter 5 presented the analysis problem of distance estimation (finding the distance of closest approach). In contrast, this chapter focuses on the control program of maximizing the minimum distance.

The distance-maximizing control problem with a point-unsafe-set distance function $c(x; X_u) = \inf_{y \in X_u} c(x, y)$ and an admissible input set \mathcal{U} (taking values in $U \subset \mathbb{R}^L$) is

$$Q^* = \sup_{u, x_0, t^*} \inf_{t \in [0, t^*]} c(x(t \mid x_0, u(\cdot)); X_u) \quad (8.1a)$$

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \forall t \in [0, t^*] \quad (8.1b)$$

$$x(0) = x_0 \in X_0, x(T \mid x_0, u(\cdot)) \in X_T \quad (8.1c)$$

$$u \in \mathcal{U}, t^* \in [0, T]. \quad (8.1d)$$

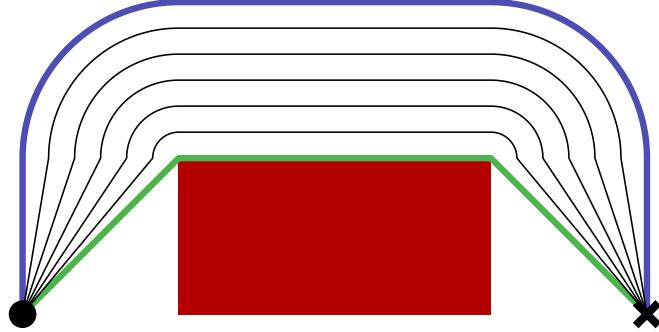
The optimization variables of (8.1) are the control policy $u(\cdot)$ and the initial point $x_0 \in X_0$.

An example of distance maximizing control is in Figure 8.1. The initial set X_0 is the black circle, and the terminal set X_T is the black ‘x’ symbol. The points X_0 and X_T are each a distance of 1.5 away from X_u . The optimum value of Q^* from (5.9) is therefore upper-bounded by 1.5, and will attain this upper bound given an arbitrarily large time horizon T . Simple integrator dynamics $\dot{x} = u$ are used for steering with $u \in [-1, 1]^2 = U$. The minimum possible time to travel from X_0 to X_T while staying at a distance of 1.5 away from X_u is $T_{arc} = 10.7124$ time units. The minimum-time control reaching an infinitesimally small distance away from X_u is $T_{min} = 7.2426$.

CHAPTER 8. DISTANCE-MAXIMIZING CONTROL

Constant-distance curves are plotted in black for time budgets $T \in (T_{\min}, T_{\text{arc}})$. As the time budget T decreases, the agent gets closer to the unsafe set and ‘cuts corners.’

T_{min}: 7.2426, T_{arc}: 10.7124



Time for Distance-Bounded Control

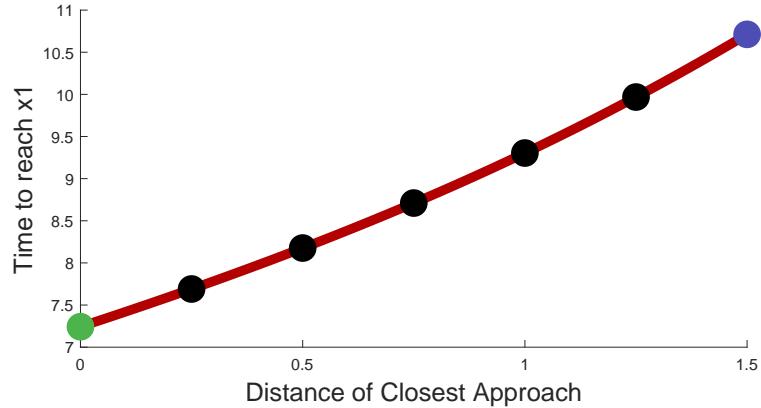


Figure 8.1: Distance maximizing control avoiding a rectangular block

Section 8.1 presents the peak-minimizing-inspired [12] formulation of problem 8.1, along with its measure and function LPs. Section 8.2 uses robust counterparts to simplify the Lie constraint when c has an SDR graph and X_u is the union of SDR sets. Section 8.3 extends the distance-maximizing control framework towards maximizing the distance of all points on a shape with respect to the unsafe set (continuing the shape-distance estimation problem of Section 5.7). Section 8.4 concludes the chapter and outlines future work.

The distance-maximizing control work in Chapter 8 is performed in collaboration with Mario Sznajer. This subject is under active development; the content of this chapter is primarily theoretical. One of the main areas of future work is performing numerical testing and experiments.

8.1 Distance Maximizing Control Program

This section provides reformulations of the distance-maximizing control program (8.1).

8.1.1 Assumptions

We posit the following assumptions:

- A1 The sets $[0, T], X, X_u, X_0, X_T, U$ are all compact and $X_0 \subset X, X_u \subset X$.
- A2 The distance function c is continuous.
- A3 The dynamics function $f(t, x, u)$ is Lipschitz in the compact domain $[0, T] \times X \times U$.
- A4 If $x(t | x_0, u) \in \partial X$ for some $t \in [0, T]$, $x_0 \in X_0$, $u \in \mathcal{U}$, then $x(t' | x_0) \notin X \forall t' \in (t, T]$.
- A5 The image $f(t, x, U)$ is convex for each fixed (t, x) .

We also define D_u as the Haussdorff distance between X and X_u with

$$D_u = \sup_{x \in X} c(x; X_u). \quad (8.2)$$

The distance D_u is finite under the compactness (A1) and continuity (A2) assumptions.

8.1.2 Lifted Program

Problem (5.9) can be reformulated into a Mayer-form OCP (only terminal cost) through the addition of a new state z using the method in [12].

Theorem 8.1.1. *The following problem has the same objective as (8.1)*

$$Q_z^* = \sup_{u(t) \in U, x_0} z \quad (8.3a)$$

$$\dot{x}(t) = f(t, x(t), u(t)), \dot{z}(t) = 0 \quad \forall t \in [0, t^*] \quad (8.3b)$$

$$z \leq c(x; X_u) \quad (8.3c)$$

$$x(0) = x_0 \in X_0, x(t^* | x_0, u(\cdot)) \in X_T, z \in [0, D_u]. \quad (8.3d)$$

Proof. This proof follows by similar logic to the proof of Theorem 7.3.1. Program (8.3) involves a maximization objective, a free terminal time, a terminal state constraint, a free initial condition, and a sign-reversed support constraint (8.3c). The state z lower-bounds the worst-case distance of closest approach $\inf_{t \in [0, t^*]} c(x(t | x_0, u); X_u)$ at all times between $t \in [0, t^*]$. \square

The LP formulation of (8.3) will involve the following support sets:

$$Z = [0, D_u] \quad (8.4a)$$

$$\Omega_z = \{(x, z) \in X \times Z \mid z \leq c(x; X_u)\} \quad (8.4b)$$

$$\Omega = \{(x, y, z) \in X \times X_u \times Z \mid z \leq c(x, y)\}. \quad (8.4c)$$

The sets in (8.4a) satisfy the projection relation of

$$\Omega_z = \pi^{xz}\Omega. \quad (8.5)$$

8.1.3 Linear Programs

The Lie derivative used in this chapter is

$$\forall v \in C^1 : \quad \mathcal{L}v(t, x, z) = \partial_t v(t, x, z) + f(t, x, w) \cdot \nabla_x v(t, x, z). \quad (8.6)$$

A measure LP of (8.3) can be constructed with an initial measure μ_0 , a terminal measure μ_t , and a relaxed occupation measure μ .

Theorem 8.1.2. *The following LP in measures will produce an upper-bound of (8.1) (with $q^* \geq Q^*$) under A2 and A4:*

$$m^* = \sup_{\mu_0, \mu_p, \mu} \langle z, \mu_T \rangle \quad (8.7a)$$

$$\mu_T = \delta_0 \otimes \mu_0 + \pi_\#^{txz} \mathcal{L}^\dagger \mu \quad (8.7b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (8.7c)$$

$$\mu \in \mathcal{M}_+([0, T] \times \Omega_z \times U) \quad (8.7d)$$

$$\mu_T \in \mathcal{M}_+([0, T] \times \Omega_z), \quad \mu_0 \in \mathcal{M}_+(X_0 \times Z). \quad (8.7e)$$

Proof. This proof will proceed by constructing a measure solution (μ_0, μ_T, μ) from every controlled trajectory satisfying the constraints of (8.1). Assume that $x_0 \in X_0$ is an initial condition, $t^* \in [0, T]$ is a stopping time, and $u(\cdot) \in \mathcal{U}$ is an admissible input in the times $[0, t^*]$ under the constraint $x(t^* \mid x_0, u(\cdot)) \in X_T$. Define $z^* = \inf_{t \in [0, t^*]} c(x(t \mid x_0, u(\cdot)); X_u)$ as the distance of closest approach along the controlled trajectory. Then the probability measures may be chosen as $\mu_0 = \delta_{x=x_0^*, z=z^*}$ and $\mu_T = \delta_{t=t^*, x=x(t|x_0, u(\cdot)), z=z^*}$, and μ can be set to the occupation measure of $t \mapsto (t, x(t \mid x_0, u(\cdot)), z^*, u(t))$ in the times $t \in [0, t^*]$. \square

Remark 8.1.1. Because z is constant along trajectories, the measure $\mu_0 \in \mathcal{M}_+(X_0 \times Z)$ will be supported in $(x_0, z) \in \Omega_z \mid x = x_0$ through the Liouville equation (8.7b) and the μ_T support in (8.7e).

Theorem 8.1.3. Under assumptions A1-A5, $m^* = Q^*$ from (8.1) and (8.7).

Proof. See the proof of (7.3.3), with modifications only for the supremizing objective. \square

Theorem 8.1.4. The dual LP of (8.7) in the auxiliary function $v(t, x, z)$ and scalar γ is

$$q^* = \inf_{\gamma \in \mathbb{R}, v} \gamma \quad (8.8a)$$

$$\gamma \geq v(0, x, z) \quad \forall X_0 \times Z \quad (8.8b)$$

$$z \leq v(t, x, z) \quad \forall (t, x, z) \in [0, T] \times \Omega_z \mid_{x \in X_T} \quad (8.8c)$$

$$\partial_t v(t, x, z) + f(t, x, u) \cdot \nabla_x v(t, x, z) \geq 0 \quad \forall (t, x, z, u) \in [0, T] \times \Omega_z \times U \quad (8.8d)$$

$$v \in C^1([0, T] \times X \times Z). \quad (8.8e)$$

In addition, strong duality with $m^* = q^*$ holds under assumptions A1-A5.

Proof. Use of OCP LP duality arguments from [105, 7, 16] will prove this theorem (just like in Theorem 7.3.4). \square

8.1.4 SOS program

We will express (8.8) in terms of the support set Ω rather than the projection Ω_z when applying the Moment-SOS hierarchy. This change in support set will occur because instances where Ω is BSA may lead to cases where the projection $\Omega_z = \pi^{xz}\Omega$ may no longer be BSA.

Corollary 2. The following LP has the same optimal value and optimizers ($\gamma \in \mathbb{R}, v$) as in (8.8):

$$q^* = \inf_{\gamma \in \mathbb{R}} \gamma \quad (8.9a)$$

$$\gamma \geq v(0, x, z) \quad \forall (x, z) \in X_0 \times Z \quad (8.9b)$$

$$z \leq v(t, x, z) \quad \forall (t, x, y, z) \in [0, T] \times \Omega \mid_{x \in X_T} \quad (8.9c)$$

$$\partial_t v(t, x, z) + f(t, x, u) \cdot \nabla_x v(t, x, z) \geq 0 \quad \forall (t, x, y, z, u) \in [0, T] \times \Omega \times U \quad (8.9d)$$

$$v \in C^1([0, T] \times X \times Z). \quad (8.9e)$$

Proof. This equivalence arises from the projection (8.5). The y variable is the coordinate on the unsafe set X_u . \square

8.1.4.1 SOS Formulation

In order to apply the moment-SOS hierarchy, we require an additional assumption:

A6 The sets X_0, X_T, X, Ω, U are all Archimedean.

Given a degree d , define the dynamics degree as $\tilde{d} = d + \lfloor \deg f/2 \rfloor$. The order- d SOS truncation of (8.9) is:

$$q_d^* = \min \gamma \in \mathbb{R}, v\gamma \quad (8.10a)$$

$$\gamma - v(0, x, z) \in \Sigma[X_0 \times Z]_d \quad (8.10b)$$

$$v(t, x, z) - z \in \Sigma[[0, T] \times X \times Z]_d \quad (8.10c)$$

$$\partial_t v(t, x, z) + f(t, x, u) \cdot \nabla_x v(t, x, z) \in \Sigma[[0, T] \times \Omega \times U]_{\tilde{d}} \quad (8.10d)$$

$$v \in \mathbb{R}[t, x, z]_{\leq 2d}. \quad (8.10e)$$

Theorem 8.1.5. Program (8.10) will converge as $\lim_{d \rightarrow \infty} q_d^* = q^*$ to the objective in (8.9) under A1-A6.

Proof. This convergence will occur by similar reasoning to the proof of Theorem 7.5.2. \square

8.1.4.2 Computational Complexity

The Lie constraint in (8.9d) has $2 + 2n + L$ variables (t, x, y, z, u) . The Gram matrix of maximal size for (8.10) occurs in constraint (8.10d), and has size $\binom{2+2n+L+\tilde{d}}{\tilde{d}}$.

This size can grow extremely quickly in (n, m, \tilde{d}) . The maximal size of the (8.10d) for $(n = 3, m = 2, \tilde{d} = 4)$ is 1001. The robust counterpart method of Chapter 6 may be applied when the dynamics f are input-affine in u and the set U is SDR. Other methods of decreasing this size includes symmetry [110], term sparsity [62], and network structure [64].

8.1.4.3 Unions of Unsafe Sets

Assume that the unsafe set X_u is the union of N_u (finite) Archimedean BSA sets:

$$X_u = \bigcup_{k=1}^{N_u} X_u^k. \quad (8.11)$$

We define support sets from (8.4a) for the union in (8.11) as

$$\Omega_u^{\text{all}} = \{(x, z) \in X \times Z \mid \forall k : z \leq c(x; X_u^k)\} \quad (8.12a)$$

$$\Omega^{\text{all}} = \{(x, \{y_k\}_{k=1}^{N_u}, z) \in X \times \bigcup_j X_u^j \times Z \mid \forall k : z \leq c(x, y_k)\}. \quad (8.12b)$$

The set Ω^{all} involves $(N_u + 1)n + 1$ variables $(x, \{y^j\}_{j=1}^{N_u}, z)$ when each unsafe set X_u^j is full-dimensional. An inordinately expensive measure LP may be built according to (8.12a) with support set Ω^{all} .

Fortunately, the support set (8.12a) possesses correlative sparsity (see Section 5.6) with respect to the groups $I_j = (x, y^j, z)$ for each $j = 1..N_u$ [61]. This type of correlative sparsity decomposes a single maximal-size block for constraint (8.10d) of size $\binom{2+m+n(N_u+1)+\tilde{d}}{\tilde{d}}$ to N_u blocks of maximal-size $\binom{2+m+2n+\tilde{d}}{\tilde{d}}$. Similar savings exist for (8.10c), converting a single maximal-size block of size $\binom{2+n(N_u+1)+d}{d}$ to N_u blocks of size $\binom{2+n+d}{d}$.

8.2 Robust Formulation for Distance-Maximizing Control

This section demonstrates how infinite-dimensional robust counterparts (from Chapter 6) can be used to simplify constraints (8.9c)-(8.9d) by eliminating the extra variable $y \in X_u$. It will be assumed in this section that the graph of $c(x, y)$ and each set X_u^k in the union (8.11) are SDR. If the set U is SDR and the dynamics f are input-affine, then the input variable u can also be eliminated by using the infinite-dimensional robust counterparts from Chapter 6.

8.2.1 L2 and Union-of-Polytope Setting

This section will focus on the simplification of constraint (8.9c). We will work with the specific case where $c(x, y) = \|x - y\|_2^2$ and X_u is the union of polytopes $X_u^k = \{x \in \mathbb{R}^n \mid \Gamma_k x \leq h_k\}$. Each polytope X_u^k has describing matrices $\Gamma_k \in \mathbb{R}^{m_k \times n}$, $h_k \in \mathbb{R}^{m_k}$ for all $k \in 1..N_u$.

Under these restrictions, the set Ω^{all} from (8.12) is

$$\Omega^{\text{all}} = \{(x, \{y_k\}, z) \in X \times (\mathbb{R}^n)^{N_u} \times Z \mid \forall k : z \leq \|x - y_k\|_2\}. \quad (8.13)$$

For each given $(x, z) \in X_T \times Z$, we define the set $W(x, z)$ as

$$W(x, z) = \{\{y_k \in \mathbb{R}^n\}_{k=1}^{N_u} \mid \forall k : z \leq \|x - y_k\|_2, y_k \in X_u^k\}. \quad (8.14)$$

The space $W(x, z)$ in (8.14) has the constraint $z \leq \|x - y_k\|_2$, which is nonconvex in y . We perform a lifting to induce convexity by adding new variables $\tau_k \geq 0$:

$$\tilde{W}(x, z) = \{(\tau_k, y_k) \in \mathbb{R}^{n+1} \mid \forall k : z \leq \tau_k, \|x - y_k\|_2 \leq \tau_k, y_k \in X_u^k\}. \quad (8.15)$$

Proposition 8.2.1. *Membership of $\{y_k\} \in W(x, z)$ is equivalent to $\forall \{\tau_k\} : \{(\tau_k, y_k)\} \in \tilde{W}(x, z)$.*

Proof. In constraint (8.15), each τ_k is bounded between $[\|x - y_k\|_2, \infty)$. Application of the $\forall\{\tau_k\}$ quantifier in the statement of this proposition leads to the lower- τ -bound constraint $z \leq \|x - y_k\|_2$ from (8.14). \square

Membership in (8.15) can be written as $\forall k \in 1..N_u$:

$$h_k - \Gamma_k y \geq 0, \quad (x - y_k, \tau_k) \in Q^n, \quad \tau_k - z \geq 0. \quad (8.16)$$

For each $k \in 1..N_u$, expression (8.16) can be written (as in the framework of (6.5)) with variables $w_k = (y_k, \tau_k)$ using the parameters

$$A_k^1 = -\Gamma_k \quad e_k^1 = h_k \quad K_k^1 = \mathbb{R}_{\geq 0}^{m_k} \quad (8.17a)$$

$$A_k^2 = \begin{bmatrix} -I_n & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \quad e_k^2 = \begin{bmatrix} x \\ 0 \end{bmatrix} \quad K_k^2 = Q^n \quad (8.17b)$$

$$A_k^3 = 1 \quad e_k^3 = -z \quad K_k^3 = \mathbb{R}_{\geq 0}, \quad (8.17c)$$

forming the equations

$$\forall k \in 1..N_u, i \in 1..3 : \quad A_k^i[y_k; \tau_k] + e_k^i \in K_k^i. \quad (8.18)$$

Constraint (8.9c) can be expressed as

$$\forall(t, x, z) \in [0, T] \times X \times Z, \{(\tau_k, y_k)\} \in \tilde{W}(x, z) : v(t, x, z) - z \geq 0 \quad (8.19)$$

8.2.2 Robust Linear Constraint

We will form a robust counterpart to (8.9c) by defining multipliers $\zeta_k(t, x, z)$, $\omega_k(t, x, z)$, $\phi_k(t, x, z)$, $\lambda_k(t, x, z)$ against the constraints in (8.17).

Lemma 8.2.2. *The robust counterpart (6.6) of (8.9c) under the L_2 distance and the union-of-polyhedra unsafe set X_u is*

$$v - z \geq \sum_{k=1}^{N_u} h_k^T \zeta_k - z\omega + x^T \phi_k \quad \forall(t, x, z) \in [0, T] \times X_T \times Z \quad (8.20a)$$

$$\sum_{k=1}^{N_u} -\Gamma_k^T \zeta_k + \omega_k + (\lambda_k - \mathbf{1}_n^T \phi_k) = \mathbf{0}_{n \times 1} \quad (8.20b)$$

$$[\zeta_k; \omega_k] \in \mathbb{R}_{\geq 0}^{m_k+1}, \quad (\phi_k, \lambda_k) \in Q^n \quad \forall k \in 1..N_u. \quad (8.20c)$$

CHAPTER 8. DISTANCE-MAXIMIZING CONTROL

Proof. The uncertain parameters (y_k, τ_k) are not present in the inequality $v - z \geq 0$ (outside of the support definition). Constraint (8.9c) can therefore be written in the context of (6.6) as

$$b_0 = v(t, x, z) - z \quad a_0 = 0 \quad (8.21a)$$

$$b_k = \mathbf{0}_{n+1 \times 1} \quad a_k = \mathbf{0}_{n+1 \times 1} \quad \forall k \in 1..N_u. \quad (8.21b)$$

The expression in (8.20) is a specific instance of (6.6) with the correspondence in (8.21) and the cones of (8.17). \square

Theorem 8.2.3. *The multipliers in (8.20c) can be chosen to be continuous functions of (t, x, z) under assumptions A1-A5.*

Proof. The distance-maximizing control problem with L_2 distance and SDR unsafe sets satisfies all conditions of Theorem 6.4.2. Specifically:

- A1' The cones $\mathbb{R}_{\geq 0}, Q^n$ in (8.17) are convex, pointed, and have nonempty interior. The non-polyhedral cone $(x - y_k, \tau_k) \in Q^n$ obeys the Slater condition with an admissible interior-cone-value of $y_k \in X_u^k, \tau_k = 1 + \|x - y_k\|_2$.
- A2' The parameter set $[0, T] \times X_T \times Z$ under assumptions A1 and A2.
- A3' The problem data (a_0, b_k) are zero, and (b_0, e) are continuous (affine) in the parameters.
- A4' The matrices (A_k) are constant and both a_\bullet and G are zero.

Therefore, a continuous choice of multipliers exists. \square

8.2.3 Robust SOS Constraint

Theorem 8.2.4. *The convergent degree- d SOS constraint formulation to (8.20) (with an auxiliary function $v \in \mathbb{R}[t, x, z]_{\leq 2d}$) is*

$$v - z \geq \sum_{k=1}^{N_u} h_k^T \zeta_k - z\omega + x^T \phi_k \in \Sigma[[0, T] \times X_T \times Z]_d \quad (8.22a)$$

$$\text{coeff}_{txz} \left(\sum_{k=1}^{N_u} -\Gamma_k^T \zeta_k + \omega_k + (\mathbf{1}_n^T (q_k - \phi_k)) \right) = \mathbf{0}_{n \times 1} \quad (8.22b)$$

$$[\zeta_k; \omega_k] \in (\Sigma[[0, T] \times X_T \times Z])^{m_k+1} \quad \forall k \in 1..N_u \quad (8.22c)$$

$$\begin{bmatrix} \sum_{i=1}^n q_{ki} & \phi_{ki} \\ \phi_{ki} & q_{ki} \end{bmatrix} \in \Sigma^2[[0, T] \times X_T \times Z] \quad \forall k \in 1..N_u, i \in 1..n. \quad (8.22d)$$

Proof. Refer to Theorem 6.5.4 regarding the SOS polynomial tightness, and to Section (6.5.3) regarding the SOS-matrix representation of the SOC constraint in (8.22d) (with $\lambda_k = \sum_{i=1}^n q_{ik}$). \square

Application of the robust counterpart drops the Gram matrix of maximal size of $\binom{2n+d+2}{d}$ in (8.10c) (when applying Correlative sparsity in the X_u -union) to $2\binom{n+d+2}{2}$ in (8.22). This Gram maximal matrix size drops to $\binom{n+d+2}{2}$ if the L_1 or L_∞ distances are used instead of the standard L_2 distance.

8.2.4 Robust Lie Constraint Discussion

This section so far has focused on the terminal constraint (8.9c). An identical process may be followed for the Lie constraint (8.9d), forming an SOS program as in (8.22) with variables $(t, x, z, u) \in [0, T] \times X \times Z \times U$. The Gram matrix size for the Lie constraint would fall from $\binom{2n+L+d+2}{2}$ to $2\binom{n+d+L+2}{2}$. When U is SDR and f is input-affine, the u variables may also be included in the robust decomposition along with $\{(y_k, \tau_k)\}_k$. The maximal-size Gram matrix for the Lie constraint would then fall to $2\binom{n+d+2}{2}$, which is the same maximal size as in the terminal constraint.

8.3 Shape-Distance Maximization

This chapter continues the shape-distance analysis formulation 5.7 in the context of distance maximizing control.

This setting will involve a shape S that is traveling according to an evolving orientation $\varphi(t)$. The shape and orientation are equipped with a body-to-global coordinate transformation function $R : S \times \Phi \rightarrow X$. The initial and final sets of orientations are Φ_0 and Φ_T respectively. The shape distance-maximizing control problem is

$$Q^* = \sup_{u, x_0, t^*} \inf_{t \in [0, t^*], s \in S} c(R(s; \varphi(t | \varphi_0, u(\cdot))), X_u) \quad (8.23a)$$

$$\dot{\varphi}(t) = f(t, x(t), u(t)) \quad \forall t \in [0, t^*] \quad (8.23b)$$

$$\varphi(0) = \varphi_0 \in \Phi_0, \varphi(T | x_0, u(\cdot)) \in \Phi_T \quad (8.23c)$$

$$u \in \mathcal{U}, t^* \in [0, T]. \quad (8.23d)$$

8.3.1 Assumptions

This section will use assumptions from 5.7 as modified for the controlled case:

CHAPTER 8. DISTANCE-MAXIMIZING CONTROL

A1' The sets $[0, T]$, Φ , Φ_0 , Φ_T , S , X , X_u , U are compact and $\Phi_0, \Phi_T \subset \Phi$.

A2' The function $f(t, \varphi, u)$ is Lipschitz in the compact $[0, T] \times \Phi \times U$.

A3' The distance $c(x, y)$ is continuous.

A4' The coordinate transformation function $R(s; \varphi)$ is continuous.

A5' If $\varphi(t \mid \varphi_0, u) \in \partial\Phi$ for some $t \in [0, T]$, $\varphi_0 \in \Phi_0$, $u \in \mathcal{U}$, then $\varphi(t \mid \varphi_0, u) \notin \Phi \forall t' \in (t, T]$.

A6' If $\exists s \in S$ such that $R(s; \varphi(t \mid \varphi_0, u)) \notin X$ or $R(s; \varphi(t \mid \varphi_0, u)) \in \partial X$ for some $t \in [0, T]$, $\varphi_0 \in \Phi_0$, $u \in \mathcal{U}$, then $R(s; \varphi(t' \mid \varphi_0, u)) \notin X \forall t' \in (t, T]$.

8.3.2 Shape Distance Maximizing Problem

Problem (8.23) can be converted into a Mayer-form OCP by adding a new state z [12]:

$$Q^* = \sup_{u, x_0, t^*} z \quad (8.24a)$$

$$\dot{\varphi}(t) = f(t, \varphi(t), u(t)), \dot{z}(t) = 0 \quad \forall t \in [0, t^*] \quad (8.24b)$$

$$z \geq c(\varphi(R(s; \varphi), \varphi_0, u); y) \quad \forall y \in X_u, s \in S \quad (8.24c)$$

$$\varphi(0) = \varphi_0 \in \Phi_0, \varphi(T \mid x_0, u(\cdot)) \in \Phi_T \quad (8.24d)$$

$$u \in \mathcal{U}, t^* \in [0, T]. \quad (8.24e)$$

The support sets of (x, z) are

$$\Omega_u^s = \{(\varphi, z) \in \Phi \times Z \mid \forall s \in S : z \leq c(R(s; \varphi); X_u)\} \quad (8.25a)$$

$$\Omega^s = \{(\varphi, s, y, z) \in \Phi \times S \times X_u \times Z \mid z \leq c(R(s; \varphi), y)\}. \quad (8.25b)$$

The LP of (8.24) involving an auxiliary function $v(t, \varphi, z)$ (similar to (8.9)) is

$$q^* = \inf_{\gamma \in \mathbb{R}} \gamma \quad (8.26a)$$

$$\gamma \geq v(0, \varphi, z) \quad \forall (\varphi, z) \in X_0 \times Z \quad (8.26b)$$

$$z \leq v(t, \varphi, z) \quad \forall (t, \varphi, s, y, z) \in [0, T] \times \Omega^s \mid_{x \in X_T} \quad (8.26c)$$

$$\partial_t v(t, \varphi, z) + f(t, \varphi, u) \cdot \nabla_x v(t, \varphi, z) \geq 0 \quad \forall (t, \varphi, s, y, z, u) \in [0, T] \times \Omega^s \times U \quad (8.26d)$$

$$v \in C^1([0, T] \times X \times Z). \quad (8.26e)$$

The dominant computational cost of an SOS tightening of (8.26) will occur with the $2 + 2n + L + N_\omega$ variables (t, φ, s, y, z, u) in (8.26d).

8.3.3 Robust Counterparts

Robust counterparts from Section 8.2 may be applied if X_u and S are both unions of SDR sets and the graph of c is SDR. The shape-modified set \tilde{W} from (8.15) with N_u unsafe sets X_u^k and N_s shapes $S^{k'}$ has variables

$$w = \{(\tau_k, y_k) \in \mathbb{R}^{n+1}\}_{k=1}^{N_u} \{s_{k'}\}_{k'=1}^{N_s}\}, \quad (8.27a)$$

leading to the constraint description of

$$\tilde{W}^s(\varphi, z) = \{w \in \mathbb{R}^{N_u(n+1)+N_s n} \mid \forall k, k' : z \leq \tau_k, \|R(s_{k'}; \varphi) - y_k\|_2 \leq \tau_k, y_k \in X_u^k, s_{k'} \in S^{k'}\}. \quad (8.27b)$$

The robust counterpart of (8.26c) and (8.26d) will be nonconservative when $R(s; \varphi)$ is an affine function, because conditions A3' and A4' from Theorem 6.4.2 will be satisfied.

In the case where $R(s; \varphi)$ involves a rigid body transformation (such as in Section 5.8.3 with the rotating square) with affine transformation $\text{Aff}(\varphi)$ and rotation $\text{Rot}(\varphi)$, the description of the R -involved constraint of (8.27b) is

$$\|\text{Rot}(\varphi)s_{k'} + \text{Aff}(\varphi) - y_k\|_2 \leq \tau_k, \quad (8.28)$$

which can be expressed in conic form from (6.5) (involving only the variables $w_{k',k} = [s_{k'}; y_k; \tau_k]$) as

$$A = \begin{bmatrix} \text{Rot}(\varphi) & -I_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix} \quad e = \begin{bmatrix} \text{Aff}(\varphi) \\ \mathbf{0} \\ 0 \end{bmatrix} \quad K = Q^n. \quad (8.29)$$

The matrix A in (8.29) depends on the parameter φ , and therefore violates assumption A4' of Theorem 6.4.2. We conjecture that for the specific rigid-body transformation case, the map Ψ_ρ from (A.10d) is Lower Semicontinuous. As a result, the main result of Theorem 6.4.2 will remain valid (allowing for the choice of continuous multiplier functions). Future work will involve proving this lower semicontinuity in the rigid body transformation case.

8.4 Conclusion

This chapter provided theory for distance-maximizing control. The distance-maximizing control problem (8.1) can be transformed into a Mayer OCP through the peak-minimizing method of

CHAPTER 8. DISTANCE-MAXIMIZING CONTROL

[12], and then solved by infinite-dimensional LPs in occupation measures [7]. However, these infinite-dimensional LP and their derived finite-dimensional SDPs suffer from a large number of variables (t, x, y, z, u) in (8.10d). Robust counterparts can be used to eliminate the unsafe-set coordinates y when the graph of c is SDR (e.g., L_2) and when X_u is the union of SDR sets. Distance-maximizing control was also extended to maximize the distance between any point on a shape and the unsafe set while steering to the destination.

Future work will involve experiments and verification. One area involves numerical optimal control. I tried and failed to use the ACADOS, CASADI [104], and GPOPS OCP solvers to approximate (8.1) numerically. The bounds of the undecomposed SOS programs (8.10) were consistently poor in simple examples (almost always D_u), and the large number of variables in (8.10d) made it difficult to raise the SDP to high degrees (Mosek/Yalmip ran out of memory).

I also ran into time constraints when preparing this thesis regarding implementation of the robust counterparts. Following the presentation of this thesis, I will implement the robust counterpart programs, and revisit numerical OCP solvers for (8.1).

One theoretical aspect of future work is changing the objective of (8.1a) to

$$Q^* = \sup_{u, t^*} \inf t \in [0, t^*], x_0 \in X_0 c(x(t | x_0, u(\cdot)); X_u). \quad (8.30)$$

The initial condition x_0 in (8.30) is chosen in an adversarial manner to form a worst-case distance of closest approach. It is of vital interest to determine whether the objective (8.30) can be expressed as part of a convex OCP LP and solved, or whether the adversarial x_0 variable would cause consistent nonconvexity.

Part 3: Peak Estimation Extensions to non-ODE Systems

Chapter 9

Peak Estimation for Hybrid Systems

9.1 Introduction

This chapter interprets and extends the peak estimation problem to dynamical systems with hybrid behavior. A hybrid system is a dynamical system that possesses both continuous-time and discrete-time dynamics [111]. Hybrid systems have a wide array of applications, including walking robots [112], power converters [113], sampled-data control [114], and systems biology [115]. In this work (extending methods from [116]), the hybrid system is defined with respect to a series of spaces known as ‘locations’ in which the hybrid trajectory evolves according to per-location ODE dynamics. When the hybrid trajectory encounters a guard surface, it will transition to a (possibly) new location according to a reset map and continue its ODE evolution. Peak estimation of hybrid systems equips each location with a state function, and the output of the peak estimation problem is the maximum state function value obtained across all locations by all hybrid systems trajectories starting from a set of initial conditions in a given time horizon.

Measures and the Moment-SOS hierarchy have been applied to solve problems featuring hybrid dynamical systems. Instances of these extensions include OCPs [116, 117, 118] and reachable sets [119, 120]. Barrier functions to certify safety of hybrid system trajectories with respect to unsafe sets may also be found by SOS programming [36].

The chapter is organized as follows. Section 9.2 introduces preliminaries about behavior and execution of hybrid systems. Section 9.3 formulates an infinite-dimensional measure program for peak estimation of hybrid system and its associated LMI relaxation. Section 9.4 extends the hybrid peak estimation framework to safety analysis and possibly uncertain dynamical systems. Numerical examples are presented in Section 9.5. The chapter is concluded in Section 9.6. This work appears in

[121] and was coauthored by Mario Sznaier.

9.2 Hybrid Systems Preliminaries

The hybrid systems in this chapter are posed over a set of L locations. Each location $\ell = 1..L$ has state variables x_ℓ contained in the space $X^\ell \subseteq \mathbb{R}^{n_\ell}$. The subsystems obey nominal locally Lipschitz dynamics f_ℓ that satisfy

$$\dot{x}_\ell(t) = f_\ell(t, x_\ell(t)) \quad \forall \ell = 1..L. \quad (9.1)$$

Available transitions between subsystems may be represented by a directed multigraph. A multigraph is a graph where pairs of vertices may be connected by multiple distinct edges [122]. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a multigraph where each of the L vertices of \mathcal{V} corresponds to a location. Each edge $e \in \mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is a directed arc from a source $\text{src}(e)$ to a destination $\text{dst}(e)$. Self-loops with $\text{src}(e) = \text{dst}(e)$ are permitted in this class of multigraphs. Edges e are associated with a guard S_e and a reset map R_e . The guard S_e is a subset of $X_{\text{src}(e)}$, and the reset map $R_e : X_{\text{src}(e)} \rightarrow X_{\text{dst}(e)}$ effects the transition. The hybrid system is fully encoded by the tuple $\mathcal{H} = (X, f, \mathcal{G}, S, R)$ with attributes:

$X = \{X_\ell\}_{\ell=1}^L$	State Spaces
$f = \{f_\ell\}_{\ell=1}^L$	Dynamics
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	Transition Multigraph
$S = \{S_e\}_{e \in \mathcal{E}}$	Guard Surfaces
$R = \{R_e\}_{e \in \mathcal{E}}$	Reset Maps

Execution of a hybrid system with multigraph transitions is based on Algorithm 1 of [119]. An additional input is a set of Zeno caps $\{N_e\}_{e \in \mathcal{E}}$ which halt trajectory execution if any transition e is traversed at least N_e times [123]. The output of the following Algorithm 2 is a system trajectory $x(t)$, as well as records \mathcal{T}, \mathcal{C} containing information about the times and locations of state transitions respectively.

The trajectory $x(t)$ is well-defined when the time horizon T and Zeno caps N_e for all $e \in \mathcal{E}$ are finite and the guard surfaces S_e are codimension-1. The trajectory $x(t)$ induces a function $\text{Loc} : [0, T] \rightarrow 1..L$ which returns the residing location of $x(t)$ at time t . Execution requires the following assumption of transversality,

Algorithm 2: Execution of Hybrid System \mathcal{H}

Input : Initial Point x_0 , Initial Location ℓ_0 , Hybrid System \mathcal{H} , Maximal Time T ,
 Transition Caps N

Output : Trajectory of System $x(t)$, Time Breaks \mathcal{T} , Location Breaks \mathcal{C} , Transition
 Counts \mathcal{N}

Initialize Trajectory $t \leftarrow 0$, $\ell \leftarrow \ell_0$, $x(0) \leftarrow x_0$

Initialize Traces $\mathcal{T} \leftarrow \{0\}$, $\mathcal{C} \leftarrow \{\ell\}$, $\mathcal{N} \leftarrow \{0\}_{e \in \mathcal{E}}$

Loop

Follow dynamics $x'(s) = f_\ell(t, x(s))$ until $x(t)$ reaches a guard or $t = T$.

if $t = T$ **OR** $\exists S_e : x(t) \in S_e$ and $\text{src}(e) = \ell$, **OR** $\exists e : \mathcal{N}_e = N_e$ **then**
 | **halt**

end

Find a guard S_e with $x(t) \in S_e$ and $\text{src}(e) = \ell$

Append t to \mathcal{T} and $\text{dst}(e)$ to \mathcal{C}

Increment $\mathcal{N}_e \leftarrow \mathcal{N}_e + 1$

Transition to $\ell \leftarrow \text{dst}(e)$, $x(t) \leftarrow R_e(x(t))$

EndLoop

A1 Let $x_\ell(t)$ be a segment of this trajectory that emerged from a transition (ℓ', ℓ) at time t^- . For all guards S^e with $\text{src}(e) = \ell$ such that $x_\ell(t) \in S^e$, the dynamics vector $f(t, x_\ell(t))$ possesses a normal component with respect to the tangent space of S^e at $x_\ell(t)$.

Remark 9.2.1. *Assumption A1 implies that the time elapsed between any two resets is bounded below by some $\delta > 0$.*

9.3 Peak estimation of hybrid systems

This section will formulate a measure LP to upper-bound (9.2).

9.3.1 Peak Program

Let $X_0 = \{X_{0\ell}\}$ be the set of initial conditions for system trajectories. Each of these system trajectories lie inside the set $X = \{X_\ell\}$.

Each location ℓ has a state cost $p_\ell : X_\ell \rightarrow \mathbb{R}$ and a set of initial conditions $X_{0\ell} \subset X_\ell$. Each p_ℓ is either bounded below or constant at $-\infty$, and at least one p_ℓ is bounded. The goal of peak estimation is to find the trajectory $x(t)$ which maximizes the state cost across all trajectories and locations

$$P^* = \sup_{t, \ell_0} \max_{x_0} p_\ell(x(t | x_0)) \quad x(t) \in X_\ell$$

Dynamics follow Algorithm 2 with input $(\ell_0, x_0, \mathcal{H}, T)$

$$x_0 \in X_{0\ell_0}. \tag{9.2}$$

The optimization variables of (9.2) are the peak time t , initial location ℓ_0 , and initial state $x_0 \in X_{\ell_0}$. The inner maximization runs over all location-objective functions p_ℓ .

The following assumptions will be posed on problem (9.2):

A1 The set $[0, T], X_\ell, X_{0\ell}$ are compact $\forall \ell = 1..L$.

A2 Problem (9.2) has a finite objective $P^* < \infty$.

A3 Each dynamics function $f_\ell(t, x_\ell)$ is Lipschitz over the compact set $[0, T] \times X_\ell$.

A4 If a trajectory at location ℓ reaches the boundary $x(t | x_0) \in \partial X_\ell$ and there does not exist a guard S_e with $x(t | x_0) \in S_e$ and $\text{src}(e) = \ell$, then $x(t | x_0)$ remains outside $\cup_\ell X_\ell$ for all $t' \in (t, T]$.

Assumption A4 is a hybrid version of the non-return assumption from Section ??.

9.3.2 Measures for Hybrid Systems

The control and reachability set programs in [119, 120, 124] define measures ρ_e supported over the guard $\mathcal{M}_+(S_e)$ for each transition $e \in \mathcal{E}$. For subsets $A \subset [0, T]$, $C_e \subset S_e$ and an initial condition x_0 , the counting measure ρ_e records the number of times the trajectory, starting from location $\text{src}(e)$, enters the patch C_e of the guard S_e with

$$\rho_e(A \times C_e) = \int_A \text{card} \left(\lim_{t' \rightarrow t^-} x(t' | x_0) \in C_e \right) dt. \quad (9.3)$$

The mass of the counting measure ρ_e is the expected number of times a trajectory will traverse the transition with arc e . In a Zeno execution of transition e , the mass $\langle 1, \rho_e \rangle$ will be unbounded, and constraints such as $\langle 1, \rho_e \rangle \leq N_e$ may be imposed to cap the maximum number of transitions on arc e . Let $X_{0\ell} \subseteq X_\ell$ be a set of initial conditions defined on each space X_ℓ in X . A distribution of initial conditions over each location is $\mu_{0\ell} \in \mathcal{M}_+(X_{0\ell})$ for $\ell = 1..L$. Let $T < \infty$ be a final time, and $\mu_{p\ell} \in \mathcal{M}_+([0, T] \times X_\ell)$ be peak measures supported over each location-space. Trajectories following dynamics $x'(t) = f_\ell(t, x(t))$ in each space X_ℓ are tracked by occupation measures $\mathcal{M}_+([0, T] \times X_\ell)$. Counting measures $\rho_e \in \mathcal{M}_+(S_e)$ are set up over all guards to handle state transitions. The Liouville equation with guard measures holding for all test functions $v_\ell \in C^1([0, T] \times X_\ell)$ and locations $\ell = 1..L$ is

$$\mu_{p\ell} = \delta_0 \otimes \mu_{0\ell} + \mathcal{L}_{f_\ell}^\dagger \mu_\ell + \sum_{\text{src}(e)=\ell} R_{e\#} \rho_e - \sum_{\text{dst}(e)=\ell} \rho_e. \quad (9.4)$$

For a location ℓ and edge e with $\text{src}(e) = \ell$, the pushforward term $R_{e\#}$ in (9.4) should be understood as

$$\langle v_\ell, R_{e\#} \rho_e \rangle = \langle v_\ell(t, R_e(x_\ell)), \rho_e \rangle. \quad (9.5)$$

The mass of the peak measure $\mu_{p\ell}$ is equal to the mass of the initial measure $\mu_{0\ell}$ plus the net flux due to state transitions.

9.3.3 Measure Program

Problem (9.2) may be relaxed through an infinite-dimensional linear program in occupation measures. The measures $\mu_{0\ell}$ are distributions of initial conditions and ρ_e are transition counting measures, just as in the Liouville equation (9.4). The peak measures $\mu_{p\ell}$ are final measures with free

CHAPTER 9. PEAK ESTIMATION FOR HYBRID SYSTEMS

terminal time between $t \in [0, T]$. The measure program in terms of $(\mu_0, \mu_p, \mu, \rho)$ for hybrid peak estimation is (where $\forall \ell$ and $\forall e$ may be expanded to $\forall \ell = 1..L$ and $\forall e \in \mathcal{E}$)

$$p^* = \sup \quad \sum_{\ell=1}^L \langle p_\ell, \mu_{p\ell} \rangle \quad (9.6a)$$

$$\mu_{p\ell} = \delta_0 \otimes \mu_{0\ell} + \mathcal{L}_{f_\ell}^\dagger \mu_\ell \quad \forall \ell \quad (9.6b)$$

$$+ \sum_{\text{dst}(e)=\ell} R_{e\#} \rho_e - \sum_{\text{src}(e)=\ell} \rho_e$$

$$\sum_{\ell=1}^L \langle 1, \mu_{0\ell} \rangle = 1 \quad (9.6c)$$

$$\langle 1, \rho_e \rangle \leq N_e \quad \forall e \quad (9.6d)$$

$$\mu_\ell, \mu_{p\ell} \in \mathcal{M}_+([0, T] \times X_\ell) \quad \forall \ell \quad (9.6e)$$

$$\mu_{0\ell} \in \mathcal{M}_+(X_{0\ell}) \quad \forall \ell \quad (9.6f)$$

$$\rho_e \in \mathcal{M}_+(S_e) \quad \forall e. \quad (9.6g)$$

Theorem 9.3.1. *Solutions to (9.6) and (9.2) satisfy $p^* \geq P^*$*

Proof. Let $(x(t \mid x_0, \ell_0), \mathcal{T}, \mathcal{C})$ be a trajectory from the execution of Algorithm 2 that stops at time $t^* \in [0, T]$, and $\text{Loc}(t)$ be the function returning the residing location of $x(t)$ at time t . This trajectory may be described by a tuple (ℓ_0, x_0, t^*) . Measures $\forall \ell : \mu_{0\ell}, \mu_{p\ell}, \mu_\ell$ and $\forall e : \rho_e$ that are feasible solutions to constraints (9.6b)-(9.6g) may be formed from the trajectory $x(t)$. The initial measure $\mu_{0\ell}$ is $\delta_{x=x_0}$ for $\ell = \ell_0$ and is the zero measure for $\ell \neq \ell_0$. The peak measure $\mu_{p\ell}$ is $\delta_{t=t^*} \otimes \delta_{x=x(t^* \mid x_0, \ell_0)}$ for $\ell = \text{Loc}(t^*)$ and is also the zero measure for all other ℓ . Let T_ℓ be the set $T_\ell = \{t \mid t \in [0, t_p^*], \text{Loc}(t) = \ell\}$ of times where $x(t)$ is in location ℓ . Each relaxed occupation measure μ_ℓ may respectively be set to the occupation measure of $t \mapsto (t, x(t \mid x_0, \ell_0))$ in the times $t \in T_\ell$. If the transition with edge $e \in \mathcal{E}$ is traversed \mathcal{N}_e times along the trajectory $x(t)$ at points $\{(t_i^e, x_i^e)\}_{i=1}^{\mathcal{N}_e}$ for $x_i^e \in X_{\text{src}(e)}$, the guard measure ρ_e may be defined as $\rho_e = \sum_{i=1}^{\mathcal{N}_e} \delta_{t=t_i^e} \otimes \delta_{x=x_i^e}$. The objective p^* is an upper bound on P^* because a set of measures $(\mu_{0\ell}, \mu_{p\ell}, \mu_\ell, \rho_e)$ constructed from every trajectory $x(t)$ satisfy the constraints of (9.6) with objective P^* . \square

Remark 9.3.1. *Setting a peak objective to $p_\ell(x) = -\infty$ is equivalent to constraining $\mu_{p\ell}$ to the zero measure, because trajectories to maximize $p(x)$ will not terminate in location ℓ . Likewise, a measure $\mu_{0\ell} \in \mathcal{M}_+(X_{0\ell})$ where $X_{0\ell} = \emptyset$ is the zero measure.*

Theorem 9.3.2. *All measures involved in a solution to (9.6) are bounded.*

Proof. Sufficient conditions for a measure to be bounded are that its mass is finite and its support is compact. This setting satisfies the compact support requirement.

Given that all measures $(\mu_0, \mu_p, \mu, \rho)$ are nonnegative, their masses will also be nonnegative numbers. The mass of the transition measures ρ are upper bounded by the Zeno constraints (9.6d) under the assumption that all N_e are finite. Constraint (9.6f) upper bounds each mass $\langle 1, \mu_{0\ell} \rangle$. For each location ℓ , choosing a test function $v_\ell(t, x_\ell) = 1$ for Liouville equation (9.6b) yields

$$\langle 1, \mu_{p\ell} \rangle = \langle 1, \mu_{0\ell} \rangle + \sum_{\text{dst}(e)=\ell} \langle 1, \rho_e \rangle - \sum_{\text{src}(e)=\ell} \langle 1, \rho_e \rangle. \quad (9.7)$$

Every term on the right-hand side of (9.7) is finite and $\langle 1, \mu_{p\ell} \rangle \geq 0$ by measure nonnegativity, so each peak measure $\mu_{p\ell}$ has bounded mass. Utilizing a test function of $v_\ell(t, x_\ell) = t$ with $\mathcal{L}_{f_\ell} t = 1$ results in

$$\langle t, \mu_{p\ell} \rangle = \langle 1, \mu_\ell \rangle + \sum_{\text{dst}(e)=\ell} \langle t, \rho_e \rangle - \sum_{\text{src}(e)=\ell} \langle t, \rho_e \rangle. \quad (9.8)$$

The terms $\langle t, \mu_{p\ell} \rangle, \langle t, \rho_e \rangle$ are all finite due to bounded masses and compact support, so the occupation measures μ_ℓ also have finite mass and are bounded. \square

Theorem 9.3.3. *The objectives in (9.2) and (9.6) will satisfy $p^* = P^*$ when $[0, T] \times \prod_\ell X_\ell$ is compact, each f_ℓ is Lipschitz, and p^* is bounded above.*

Proof. This statement may be proved by extending arguments from [116]. Theorem 17 of [116] states there is no relaxation gap in measure LPs of an optimal control program with appropriate assumptions, extending the ODE result of [7]. Free final time is already accounted for in [116] by reference to Remark 2.1 of [16]. The ODE problem in [7] can handle initial conditions lying in a set X_0 , so the method in [116] can similarly work with sets of initial conditions $\{X_{0\ell}\}_{\ell=1}^L$ as demonstrated by [117]. The work in [117] has ‘switching’ costs (possibly differing running and terminal costs in each location), which is realized by the costs p_ℓ . The final modification between this work and [116] is that problem (9.2) has finite Zeno caps N_e , while Assumption 3 of [116] forbids Zeno trajectories. The allowance for free terminal time permits consequence 4 of Theorem 12 of [116] to read that there exists a constant C such that $\sum_e \langle 1, \rho_e \rangle \leq \sum_e N_e = C$. The three modifications of [116] (free terminal time, multiple initial conditions, Zeno caps) are all cleared, so $p^* = P^*$ under the compactness and Lipschitz assumptions. \square

9.3.4 Function Program

The measure program (9.6) is dual to an infinite-dimensional linear program in continuous functions. The Lagrangian \mathcal{L} of problem (9.6) with dual variables $v_\ell \in C^1([0, T] \times X_\ell)$, $\gamma \in$

CHAPTER 9. PEAK ESTIMATION FOR HYBRID SYSTEMS

$\mathbb{R}, \alpha \in \mathbb{R}_+^{|\mathcal{E}|}$ is

$$\begin{aligned}\mathcal{L} = & \sum_{\ell=1}^L \langle p_\ell, \mu_{p\ell} \rangle + \langle v_\ell(t, x), \delta_0 \otimes \mu_{0\ell} + \mathcal{L}_{f_\ell}^\dagger \mu_\ell \rangle \\ & + \langle v_\ell(t, x), \sum_{\text{dst}(e)=\ell} R_e \# \rho_e - \sum_{\text{src}(e)=\ell} \rho_e - \mu_{p\ell} \rangle \\ & + \gamma (1 - \sum_{\ell=1}^L \langle 1, \mu_{0\ell} \rangle) + \sum_{e \in \mathcal{E}} \alpha_e (N_e - \langle 1, \rho_e \rangle).\end{aligned}\quad (9.9)$$

The dual function program of (9.6) is

$$\begin{aligned}d^* = & \inf_{\gamma, \alpha, v} \sup_{\mu_{0\ell}, \mu_{p\ell}, \mu_\ell, \rho_e} \mathcal{L} \\ d^* = & \inf_{\gamma \in \mathbb{R}, \alpha \in \mathbb{R}_+^{|\mathcal{E}|}} \gamma + \sum_{e \in \mathcal{E}} N_e \alpha_e\end{aligned}\quad (9.10a)$$

$$\forall \ell : \forall x_\ell \in X_{0\ell} :$$

$$\gamma \geq v_\ell(0, x_\ell) \quad (9.10b)$$

$$\forall \ell : \forall (t, x_\ell) \in [0, T] \times X_\ell :$$

$$0 \geq \mathcal{L}_{f_\ell} v_\ell(t, x_\ell) \quad (9.10c)$$

$$\forall e : \forall (t, x_{\text{src}(e)}) \in [0, T] \times X_{\text{src}(e)} :$$

$$v_{\text{src}(e)}(t, x_{\text{src}(e)}) - v_{\text{dst}(e)}(t, R_e(x_{\text{src}(e)})) \geq -\alpha_e \quad (9.10d)$$

$$\forall \ell : \forall (t, x_\ell) \in [0, T] \times X_\ell :$$

$$v_\ell(t, x_\ell) \geq p_\ell(x_\ell) \quad (9.10e)$$

$$\forall \ell : v_\ell(t, x_\ell) \in C^1([0, T] \times X_\ell). \quad (9.10f)$$

The dual variables v_ℓ are auxiliary functions that decrease along trajectories (9.10c) and along transitions (9.10d). The auxiliary functions upper bound the location-costs by (9.10e). The dual variable α_e will be zero if transition e is traveled at most $N_e - 1$ times (complementary slackness of (9.6g)).

Theorem 9.3.4. *Programs (9.6) and (9.10) will possess equal objectives $p^* = d^*$ when each X_ℓ is compact and (T, N_e) are each finite.*

Proof. $p^* = d^*$: Strong duality follows by arguments from Theorem 2.6 of [23], specifically from boundedness of measures (Theorem 9.3.2) and compactness (Assumption A2). \square

9.3.5 Linear Matrix Inequality

The BSA sets containing measures in (9.6) are

$$\begin{aligned} \forall \ell : \quad X_\ell &= \{x_\ell \mid g_{\ell i}(x_\ell) \geq 0 \mid i = 1..N_c^\ell\} \\ \forall \ell : \quad X_{0\ell} &= \{x_\ell \mid g_{0\ell i}(x_\ell) \geq 0 \mid i = 1..N_c^{0\ell}\} \\ \forall e : \quad S_e &= \{x_{\text{src}(e)} \mid g_{ei}(x_{\text{src}(e)}) \geq 0 \mid i = 1..N_c^e\}. \end{aligned} \quad (9.11)$$

Polynomials $g_{\ell i}(x)$, $g_{0\ell i}(x_\ell)$, $g_{ei}(x_{\text{src}(e)})$ have finite degrees $d_{\ell i}$, $d_{0\ell i}$, d_{ei} respectively for each i, ℓ, e as appropriate. Let $(\mathbf{m}^{0\ell}, \mathbf{m}^{p\ell}, \mathbf{m}^\ell, \mathbf{r}^e)$ be moment sequences of the measures $(\mu_{0\ell}, \mu_{p\ell}, \mu_p, \rho_e)$. The Liouville equation (9.6b) may be expressed as a collection of affine constraints in the moment sequences. Substituting the test function $v(t, x_\ell) = x_\ell^\alpha t^\beta$ into (9.6b) yields a relation for each $\alpha \in \mathbb{N}^{n\ell}$, $\beta \in \mathbb{N}$, $\ell \in 1..L$:

$$\begin{aligned} 0 &= -\langle x_\ell^\alpha t^\beta, \mu_{p\ell} \rangle + \langle x_\ell^\alpha t^\beta, \delta_{t=0} \otimes \mu_{0\ell} \rangle + \langle \mathcal{L}_{f_\ell} x_\ell^\alpha t^\beta, \mu_\ell \rangle \\ &\quad + \sum_{\text{dst}(e)=\ell} \langle R_e(x_\ell)^\alpha t^\beta, \rho_e \rangle - \sum_{\text{src}(e)=\ell} \langle x_\ell^\alpha t^\beta, \rho_e \rangle. \end{aligned} \quad (9.12)$$

The expression $\text{Liou}_{\alpha\beta}^\ell(\mathbf{m}^{0\ell}, \mathbf{m}^{p\ell}, \mathbf{m}^\ell, \mathbf{r}^{\mathcal{E}_\ell}) = 0$ is defined to abbreviate the affine constraint in moment sequences induced by (9.12), where $\mathcal{E}_\ell = \{e \in \mathcal{E} \mid \text{src}(e) = \ell \text{ or } \text{dst}(e) = \ell\}$ is the set of arcs including location ℓ . For a constant degree $d \in \mathbb{N}$, define the quantities $d'_\ell = d + \lceil \deg f_\ell / 2 \rceil - 1$ and $k_e = \deg R_e$. The degree- d LMI relaxation of (9.6) with variables $(\mathbf{m}^{0\ell}, \mathbf{m}^{p\ell}, \mathbf{m}^\ell, \mathbf{r}^e)$ is

$$p_d^* = \max \quad \sum_\ell \sum_\alpha p_{\ell\alpha} \mathbf{m}_\alpha^{p\ell} \quad (9.13a)$$

$$\sum_\ell \mathbf{m}_0^{0\ell} = 1 \quad (9.13b)$$

$$\forall \ell : \alpha \in \mathbb{N}^{n\ell}, \beta \in \mathbb{N}, |\alpha| + |\beta| \leq 2d$$

$$\text{Liou}_{\alpha\beta}^\ell(\mathbf{m}^{0\ell}, \mathbf{m}^{p\ell}, \mathbf{m}^\ell, \mathbf{r}^{\mathcal{E}_\ell}) = 0 \text{ by (9.12)} \quad (9.13c)$$

$$\forall e : \mathbf{m}_0^e \leq N_e \quad (9.13d)$$

$$\forall \ell : \mathbb{M}_d(X^{0\ell} \mathbf{m}^{0\ell}), \mathbb{M}_d([(0, T] \times X^\ell) \mathbf{m}^{p\ell}), \mathbb{M}_{d'_\ell}([0, T] \times X^\ell \mathbf{m}^\ell) \succeq 0 \quad (9.13e)$$

$$\forall e : \mathbb{M}_{k_e d_e}(S_e \mathbf{r}^e) \succeq 0. \quad (9.13f)$$

The affine constraints (9.13c)-(9.13d) implement a truncation of constraints (9.6d)-(9.6d) in terms of finite-length moment sequences. Constraints (9.13e)-(9.13f) ensure that there exist representing measures for the moment sequences. Solutions to the SDP generated from the LMI (9.13) by raising the degree d will form a chain of upper bounds $p_d^* \geq p_{d+1}^* \geq \dots \geq p^*$.

CHAPTER 9. PEAK ESTIMATION FOR HYBRID SYSTEMS

Theorem 9.3.5. *The sequence of upper bounds will satisfy $\lim_{d \rightarrow \infty} p_d^* = P^*$ when $[0, T] \times \prod_{\ell=1}^L X_\ell$ is Archimedean, $\forall \ell : f_\ell(t, x)$ are polynomial, and $\forall e : N_e$ are finite.*

Proof. The upper bound sequence will converge to p^* when all sets are Archimedean, there exists an interior point to constraints (9.6b)-(9.6g), and all measures $(\mu_{0\ell}, \mu_{p\ell}, \mu_\ell, \rho_e)$ have bounded moments (Theorem 5 of [22] and Theorem 4.4 of [27]).

Let x_0 be an initial point starting in some nonempty location X_ℓ . The set of measures where $\mu_{0\ell} = \delta_{x=x_0}$, $\mu_{p\ell} = \delta_{t=0} \otimes \delta_{x=x_0}$ and all other measures are the zero measure is an interior point to (9.6b)-(9.6g) (trajectory starting at x_0 with zero elapsed time). Given that each $[0, T] \times X_\ell$ is compact, it is sufficient that all measures have bounded masses in order for the measures to have bounded moments. The masses of ρ_e are each upper bounded by the finite quantity N_e by constraint (9.6g), and the sum of the masses of $\mu_{0\ell}$ are upper bounded by 1 through (9.6c). The sum of constraint (9.6b) with test function $v_\ell = 1$ along all ℓ is $\sum_{\ell=1}^L \langle 1, \mu_{p\ell} \rangle = \sum_{\ell=1}^L \langle 1, \mu_{0\ell} \rangle = 1$, so each mass of $\mu_{p\ell}$ is finite. Lastly, the use of a test function of $v_\ell = t$ on each Liouville equation in (9.6b) yields the finite expression $\langle 1, \mu_\ell \rangle = \langle t, \mu_p \rangle - \sum_{\text{dst}(e)=\ell} \langle t, \rho_e \rangle + \sum_{\text{src}(e)=\ell} \langle t, \rho_e \rangle$. The sequence of upper bounds will therefore converge to p^* as $d \rightarrow \infty$ with $p^* = P^*$ from Theorem 9.3.3. \square

The sizes of the moment matrices in problem (9.13) are listed in Table 9.1. The computational complexity of numerical LMI solvers scale in a polynomial manner with the size of the largest PSD matrix [16]. These PSD matrix sizes may be reduced if extant structure such as symmetry, quotient, or sparsity structure is present in (9.6).

Table 9.1: Sizes of moment matrices in LMI (9.13)

Moment	$\mathbb{M}_d(\mathbf{m}^{0\ell})$	$\mathbb{M}_d(\mathbf{m}^{p\ell})$	$\mathbb{M}_{d'_\ell}(\mathbf{m}^\ell)$	$\mathbb{M}_{dk_e}(\mathbf{r}^e)$
Size	$\binom{n_\ell+d}{d}$	$\binom{1+n_\ell+d}{d}$	$\binom{1+n_\ell+d'_\ell}{d'_\ell}$	$\binom{1+n_\ell+k_e d}{k_e d}$

Remark 9.3.2. *Guards with codimension-1 sets S_e may replace their PSD localizing constraints with linear equality constraints or quotient ring reductions in \mathbf{r}^e .*

Remark 9.3.3. *Algorithm 1 may be used to attempt extraction of near-optimal trajectories if the moment matrices $\forall \ell : \mathbb{M}_d(\mathbf{m}^{0\ell}), \mathbb{M}_d(\mathbf{m}^{p\ell})$ are low-rank.*

9.4 Extensions

This section details extensions to the previously presented peak estimation framework for hybrid systems.

9.4.1 Uncertainty

Peak estimation for hybrid systems may be applied to systems with uncertainty, extending the ODE case from Chapter 4. Let $W_\ell \subset \mathbb{R}^{N_{w\ell}}$ be a compact set of time-dependent disturbances for each location. Each location obeys dynamics $\dot{x}_\ell = f(t, x_\ell(t), w_\ell(t))$, $\forall t, \ell : w_\ell(t) \in W_\ell$, in which there is no prior assumption of continuity on the process $w(\cdot)$. The uncertainty act as adversarial optimal controls attempting to raise the peak functions (p_ℓ).

Uncertainty in this manner may be realized by adjusting the Liouville equation in (9.6b) and occupation measure definitions in (9.6e) (where $w_\ell(t)$ acts as a Young measure/relaxed control [45]) with

$$\mu_{p\ell} = \delta_0 \otimes \mu_{0\ell} + \pi_{\#}^{tx} \mathcal{L}_{f_\ell}^\dagger \mu_\ell \quad \forall \ell \quad (9.14a)$$

$$+ \sum_{\text{dst}(e)=\ell} R_e \# \rho_e - \sum_{\text{src}(e)=\ell} \rho_e \\ \mu_\ell \in \mathcal{M}_+([0, T] \times X_\ell \times W_\ell) \quad \forall \ell. \quad (9.14b)$$

A particular form of time-dependent uncertainty is switching/polytopic structure. If the system model is $\dot{x}_\ell = \sum_k^{N_s} w_{k\ell} f_{k\ell}(t, x)$ for N_s switching modes and $w_{k\ell} \geq 0$, $\sum_k w_{k\ell} = 1$, then the Liouville equation in (9.14) may be expressed for occupation measures $\mu_{k\ell} \in \mathcal{M}_+([0, T] \times X)$ as

$$\mu_{p\ell} = \delta_0 \otimes \mu_{0\ell} + \sum_k^{N_s} \mathcal{L}_{f_{k\ell}}^\dagger \mu_{k\ell} \quad \forall \ell \quad (9.15a)$$

$$+ \sum_{\text{dst}(e)=\ell} R_e \# \rho_e - \sum_{\text{src}(e)=\ell} \rho_e \\ \mu_\ell \in \mathcal{M}_+([0, T] \times X_\ell) \quad \forall \ell. \quad (9.15b)$$

Time-independent uncertainty restricted to a compact set $\Theta \subset \mathbb{R}^{N_\theta}$ may also be added by adjoining to dynamics a new state $\dot{x}_\ell = f(t, x_\ell(t), \theta, w_\ell(t))$, $\dot{\theta} = 0$. This new state θ is preserved between transition jumps, inducing lifted reset maps $\tilde{R}_{s \rightarrow t}(x_s, \theta) = (R(x_t), \theta)$.

9.4.2 Safety

This section verifies safety of hybrid system trajectories with respect to a group of unsafe sets based on prior (ODE) work in Chapter 5. Let $X_{u\ell} = \{x_\ell \mid p_{\ell i}(x_\ell) \geq 0, i = 1..N_u\}$ be an unsafe basic semialgebraic set for each location $\ell = 1..L$.

Letting $c_\ell(x_\ell, y_\ell)$ be a distance function (cost) with a point-unsafe-set distance $c_\ell(x_\ell; X_{u\ell}) = \min_{y_\ell \in X_{u\ell}} c_\ell(x_\ell, y_\ell)$, the distance estimation problem for hybrid systems is

$$Q^* = \inf_{\ell' \in 1..N_u, t \in [0, T], \ell_0, x_0} c_\ell(x(t \mid x_0); X_{u\ell'})$$

Dynamics follow Algorithm 2 with input $(\ell_0, x_0, \mathcal{H}, T)$

$$x_0 \in X_{0\ell_0}.$$

Following the procedure from [13], a joint-measure $\eta_\ell(x_\ell, y_\ell) \in \mathcal{M}_+(X_\ell \times X_{u\ell})$ is added for each unsafe set. The distance objective in (9.16) is replaced with an equivalent expectation over the joint probability measure $\langle c_\ell(x_\ell, y_\ell), \eta_\ell \rangle$.

The measure program for distance estimation with variables $(\mu_{p\ell}, \mu_{0\ell}, \mu_\ell, \eta_\ell, \rho_e)$ is

$$q^* = \inf \sum_{\ell=1}^L \langle c_\ell(x_\ell, y_\ell), \eta_\ell \rangle \quad (9.16a)$$

$$\pi_\#^{x_\ell} \eta_\ell = \pi_\#^{x_\ell} \mu_{p\ell} \forall \ell \quad (9.16b)$$

$$\text{Constraints (9.6b)-(9.6d)} \quad (9.16c)$$

$$\text{Variables from (9.6e)-(9.6g)} \quad (9.16d)$$

$$\forall \ell : \eta_\ell(x_\ell, y_\ell) \in \mathcal{M}_+(X_\ell \times X_{u\ell}). \quad (9.16e)$$

9.5 Numerical Examples

Experiments are available at https://github.com/Jarmill/hybrid_peak_est, and were written in MATLAB 2021a. Dependencies include Gloptipoly3 [30], Yalmip interface [68], and Mosek [49]. All experiments were run on an 2.30 GHz Intel i9 CPU with 64.0 GB of RAM.

9.5.1 Two-Mode

This system is a modification of Example 2 of [36] to ensure improved numerical conditioning. The two locations correspond to modes of ‘No Control’ ($\ell = 1$) and ‘Control’ ($\ell = 2$) with

dynamics

$$\begin{aligned} f_1(t, x) &= [x_2; -x_1 + x_3; x_1 + (1 + x_3)^2(2x_2 + 3x_3)] \\ f_2(t, x) &= [x_2; -x_1 + x_3; -x_1 - (2x_2 + 3x_3)]. \end{aligned} \quad (9.17a)$$

9.5.1.1 Two-Mode: Standard

Trajectories start in the initial set $X_{01} = \{x \mid \|x\|_2^2 = 0.2^2\}$ ($X_{02} = \emptyset$), and evolve for a time horizon of $T = 20$. The transition edges are $\mathcal{E} = \{(1, 2), (2, 1)\}$ with guard surfaces

$$\begin{aligned} S_{(1,2)} &= \{x \mid x_1^2/4 + x_2^2 + x_3^2 = 1.5^2\} \\ S_{(2,1)} &= \{x \mid x_1^2 + x_2^2 + x_3^2 = 0.2^2\}. \end{aligned} \quad (9.18)$$

Each transition has a trivial reset map $R_e(x) = x$. The Zeno caps used in simulation were $N_{(1,2)} = N_{(2,1)} = 5$ with total spaces of $X_1 = X_2 = [-1.5, 1.5]^3$. Figure 9.1 plots system trajectories in location 1 (left) and 2 (right), starting from the initial set X_0 (gray region). The peak estimation task for this system is to upper bound extreme values of $p_2(x) = x_1^2$ along system trajectories ($p_1(x) = -\infty$). Solving the LMI (9.13) at orders 1-5 produces the sequence of upper bounds, $p_{1:5}^* = [2.250, 0.6514, 0.4643, 0.4076, 0.3958]$.

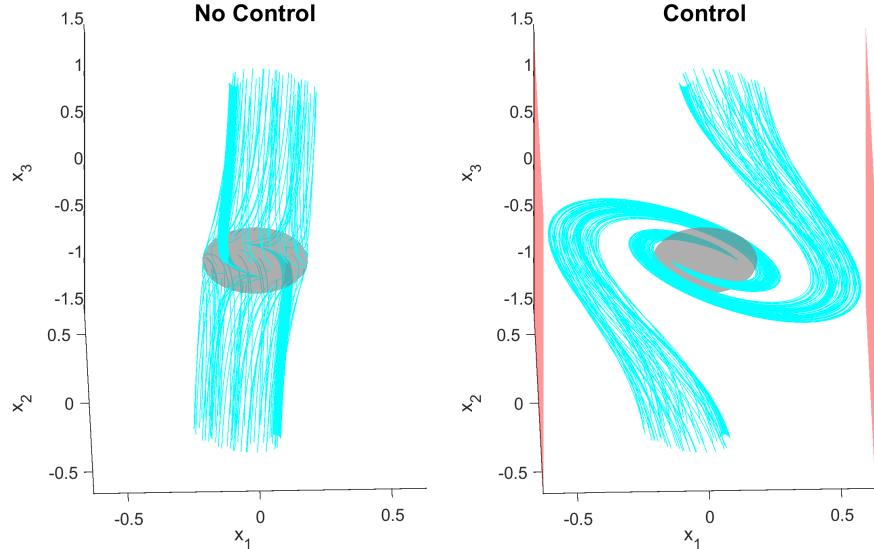


Figure 9.1: Deterministic Two-Mode Bound of $x_1^2 \leq 0.3958 = p_5^*$

9.5.1.2 Two-Mode: Uncertainty

Time-dependent uncertainty may be added to dynamics in (9.17) by defining a process $w(t) \in [-1, 1]$ under the dynamics $\tilde{f}_\ell(t, x) = f_\ell(t, x) + [0; 0; w]$. The LMI bounds for x_1^2 when w is realized as switching-type uncertainty is $p_{1:5}^* = [2.250, 1.4029, 1.0350, 0.9790, 0.9660]$. System trajectories and the order-5 bound of this noisy system are plotted in Figure 9.2.

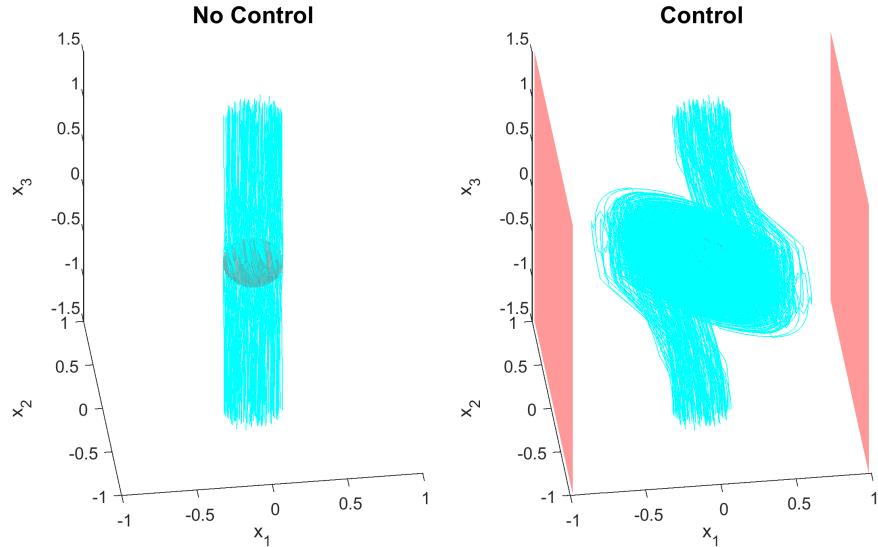


Figure 9.2: Noisy Two-Mode Bound of $x_1^2 \leq 0.9660 = p_5^*$

9.5.1.3 Two-Mode: Distance Estimation

Distance estimation is conducted for the deterministic two-mode system (9.17) with respect to the half-sphere unsafe set

$$X_u = \{x \mid 0.4^2 \geq (x_1 + 0.5)^2 + (x_2 + 0.5)^2 + (x_3 - 0.5)^2, x_3 \geq 0.5\}. \quad (9.19)$$

The distance penalty $c(x, y) = \|x - y\|_2^2$ is used in locations $\ell = 1, 2$ with the unsafe set X_u . LMI lower bounds for the distance $\min_\ell \min_{y \in X_u} \|x_\ell - y\|_2^2$ (via (9.16)) are $p_{1:5}^* = [0, 0, 0, 2.799 \times 10^{-3}, 7.942 \times 10^{-3}]$. The output of distance estimation is plotted in Figure 9.3. The solid red half-sphere is the set X_u , and the corona surrounding X_u is the set of all points with an L_2 distance at most $0.0891 = \sqrt{p_5^*}$ away from X_u .

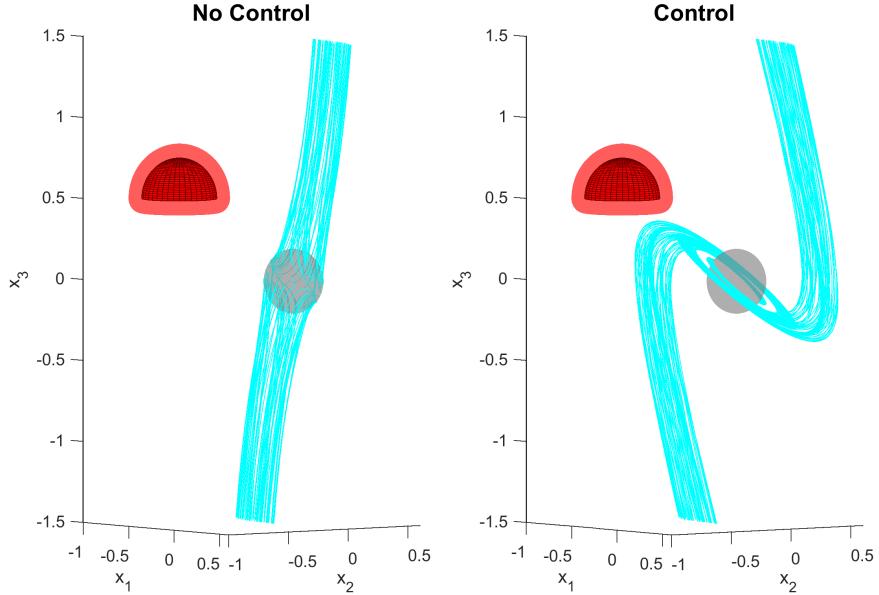


Figure 9.3: Deterministic Two-Mode Distance Bound of $\min_{y \in X_u} \|x - y\|_2 \leq 0.0891 = \sqrt{p_5^*}$

9.5.2 Right-Left Wrap

This example has a single location $X = [-1, 1]^2$ with nontrivial reset maps. The dynamics in the single location are

$$\dot{x} = \begin{bmatrix} -x_2 + x_1 x_2 + 0.5 \\ -x_2 - x_1 + x_1^3 \end{bmatrix}. \quad (9.20)$$

System (9.20) has a stable equilibrium point at $(-0.8128, 0.2758)$ and a saddle point at $(-0.4288, 0.3499)$. The following right\$\rightarrow\$top and left\$\rightarrow\$bottom transitions are defined with Zeno caps of $N = 5$

$$\begin{aligned} S_{\text{right}\rightarrow\text{top}} &= \{x \mid x_1 = 1, x_2 \in [-1, 1]\} & R_{\text{right}\rightarrow\text{top}} &= [x_2, x_1]^T \\ S_{\text{left}\rightarrow\text{bottom}} &= \{x \mid x_1 = -1, x_2 \in [-1, 1]\} & R_{\text{left}\rightarrow\text{bottom}} &= [1 - x_2, x_1]^T. \end{aligned} \quad (9.21)$$

The set X is invariant under these state transitions.

9.5.2.1 Right-Left: Standard

A peak estimation task to maximize $p(x) = -(x_1+0.5)^2 + (x_2+0.5)^2$ is defined on system dynamics (9.20) starting from the initial set $X_0 = \{x \mid 0.2^2 = (x_1-0.5)^2 + (x_2+0.3)^2\}$ for a $T = 5$ time horizon. LMI upper bounds for this objective are $p_{1:6}^* = [0, 0, -0.3644, -0.5259, -0.5659, -0.5721]$.

Figure 9.4 plots the ODE system dynamics in (9.20). Figure 9.4b plots hybrid system dynamics in cyan, starting from the black-circle X_0 . The p_6^* bound is indicated in the red circle of radius $\sqrt{-p_6^*} = 0.7564$, in which no considered hybrid system trajectory is contained.

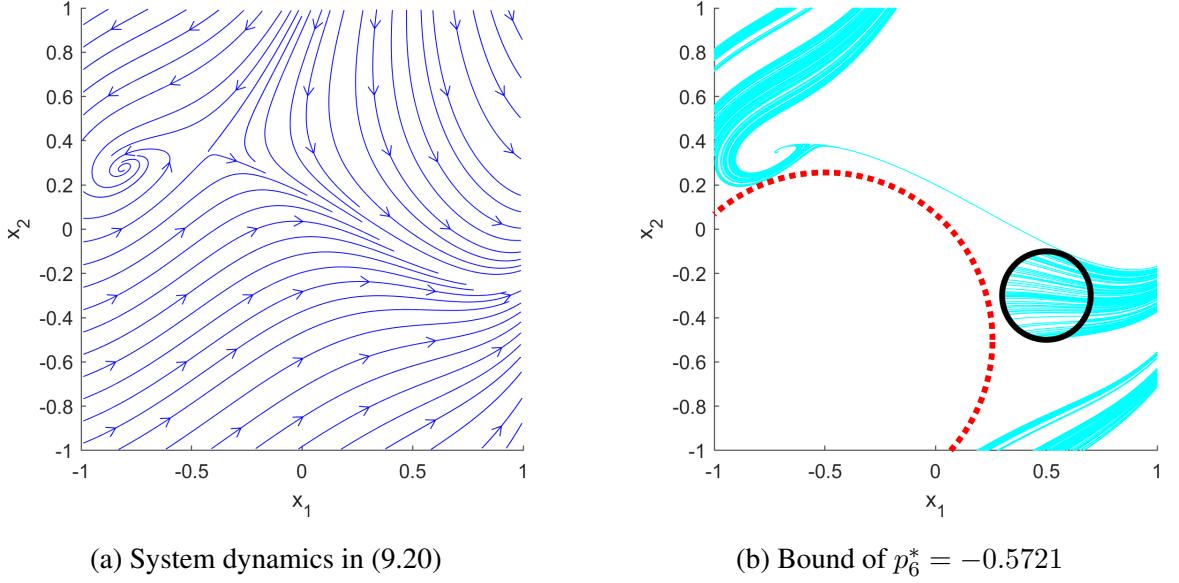


Figure 9.4: Peak Estimation of Right-Left Wrap Dynamics (9.20) and (9.21)

9.5.2.2 Right-Left: Distance

This example involves the Right-Left system (9.20) under the same X, X_0, T parameters as in Section 9.5.2.1. The half-circle unsafe set is $X_u = \{x \mid 0.4^2 \geq (x_1 - 0.5)^2 + (x_2 + 0.3)^2, (-0.6 - x_1 - x_2)/\sqrt{2} \geq 0\}$. Using a distance penalty of $c(x, y) = \|x - y\|_2^2$ leads to L_2 (square root) lower-bounds of $c_{1:6}^* = [0, 0, 0.1647, 0.2643, 0.2853, 0.2877]$. Figure 9.5 pictures the set X_u in the red half-circle, and the 0.2877-distance contour from the degree-6 SDP as the red points surrounding the half-circle.

9.6 Conclusion

An existing peak estimation framework for ODE systems (discussed in Chapter 3) was extended in this chapter to hybrid systems. A hierarchy of LMIs result in a (convergent) decreasing sequence of upper bounds to the true peak value. Extensions to the hybrid peak estimation framework, such as bounding the distance to unsafe sets (Chapter 5) and estimation of systems with uncertainty

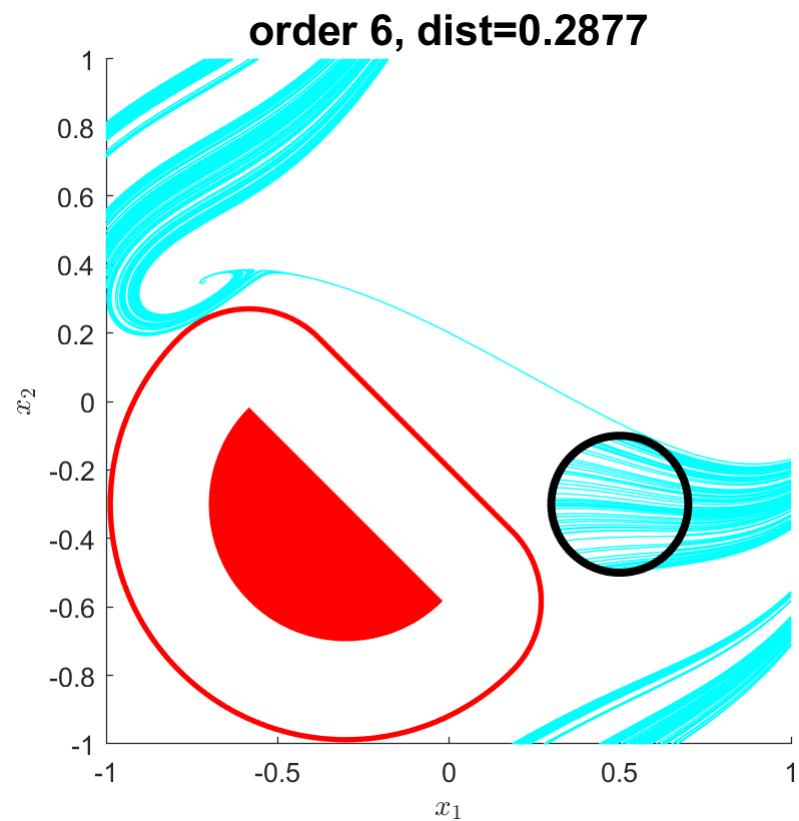


Figure 9.5: Distance bound of $\inf_{y \in X_u} \|x - y\|_2 \leq 0.2877$

CHAPTER 9. PEAK ESTIMATION FOR HYBRID SYSTEMS

(Chapter 4), can be accomplished by modifying equations in the LP. Future work includes performing peak-minimizing (L_1 -optimal) control of hybrid systems and applying numerical techniques to perform peak estimation on more complicated systems (e.g., rigid body dynamics in robotics), and generalizing analysis of deterministic hybrid dynamics to Markov Decision Processes.

Chapter 10

Value-at-Risk Peak Estimation

10.1 Introduction

This chapter analyzes maximal $(1 - \epsilon)$ -quantile statistics of a state function $p(x)$ for Stochastic Differential Equation (SDE) trajectories evolving in a compact set X . An example of this type of quantile statistic for trajectory analysis is in establishing that there exists at least one time with a 1% chance of the aircraft exceeding a height of 100 meters. This task of quantile estimation is related to peak and value-at-risk estimation, and will also be referred to as the ‘chance-peak’ problem.

The ϵ -Value-at-Risk (VaR) is the value at which there is an ϵ -probability of exceedance [125]. Control and portfolio design typically aims to minimise the VaR. One specific VaR-upper-bounding coherent risk measure [126] that results in convex programs is the conditional VaR risk measure [127, 128]. The conditional VaR has been utilized for stochastic optimal control in [129], and for approximation of discrete-time risk-bounded sets using exponential and logarithmic inequalities with Markov Decision Processes in [130]. In contrast, the chance-peak approach upper-bounds maximum VaR of the continuous-time SDE state distribution of $x(t)$ across all times. We will solve this problem by maximizing the Cantelli and Vysochanskij-Petunin (VP) upper bounds for the VaR.

Chance constraints are an adjacent topic to VaR optimization in which a probability inequality must hold as a hard constraint. Chance-constrained programs have a wide variety of application in control theory [131, 132, 74], and are generally intractable to solve explicitly. Approximation methods for chance constraints include the Cantelli [133] and VP [134] inequalities, and application of these tail-bounds in control include [135, 136]. The scenario approach for randomized constraint generation will converge in probability to the chance-constrained optimum, but carries a risk of

failure and may require a large number of samples [137]. The moment-SOS hierarchy of SDPs will converge to the chance-constrained optimal solution under appropriate boundedness conditions [138].

The chance-peak problem is also related to a family of optimal stopping problems which can be solved using occupation measures. The peak estimation work in [5] (reviewed in Chapter 3) also accounts for the stochastic case, in which finding the expectation of $p(x)$ in time is maximized. Dynamics are phrased in terms of their infinitesimal generator (Feller process) in order to pose the LPs in [5]. Such LPs will converge to the true solution of the stopping problem under mild convergence, regularity, and well-posedness assumptions. The moment-SOS hierarchy of finite-dimensional SDPs will converge to the infinite-dimensional LP optimum if all problem data (e.g., dynamics, constraint sets) are polynomial-representable [27]. Instances of the moment-SOS hierarchy being used to solve stochastic problems include expectation-maximization of Lévy processes [139], option pricing [140], [141], infinite-time averages [142], and Reach-Avoid sets [143].

This chapter has the following structure: Section 10.2 gives an overview of SDEs and occupation measures. Section 10.3 proposes an infinite-dimensional Second-Order Cone Program (SOCP) to upper-bound (10.11) in terms of occupation measures and analyzes its strong duality properties. Section 10.4 reviews the moment-SOS hierarchy and presents a hierarchy of SDPs that approximate the infinite-dimensional chance-peak SOCP. Section 10.5 details extensions to the chance-peak framework: analysis of exit-time statistics, distance of closest approach, and switching processes. Section 10.6 provides numerical examples of the chance-peak problem on ODE and SDE systems. Section 10.7 concludes the chapter. Appendix A.7 proves strong duality properties for a class of measure programs with linked semidefinite constraints. Appendix A.8 applies this general strong duality proof to the chance-peak SOCP.

This work appeared in [144] and is coauthored by Matteo Tacchi, Mario Sznaier, and Ashkan Jasour. The content of Appendix A.7 (generalized strong duality) is solely authored by Matteo Tacchi.

10.2 Preliminaries

This section will review VaR and SDEs.

10.2.1 Probability Tail Bounds and Value-at-Risk

Let ξ be a univariate probability measure $\xi(\omega) \in \mathcal{M}_+(\mathbb{R})$ for a coordinate $\omega \in \mathbb{R}$, with $\langle 1, \xi \rangle = 1$ and $|\langle \omega, \xi \rangle|, \langle \omega^2, \xi \rangle < \infty$ (finite first and second moments). In this chapter, we define the ϵ -VaR of ξ as follows:

$$VaR_\epsilon(\xi) = \sup \{ \lambda \in \mathbb{R} \mid \xi([\lambda, \infty)) \geq \epsilon \}. \quad (10.1)$$

Let $\sigma^2 = \langle \omega^2, \xi \rangle - \langle \omega, \xi \rangle^2$ be the variance of the probability distribution ξ .

The Cantelli bound for VaR is [133]

$$VaR_\epsilon(\xi) \leq \sigma \sqrt{1/(\epsilon) - 1} + \langle \omega, \xi \rangle = VaR_\epsilon^{cant}(\xi). \quad (10.2a)$$

The VP bound for the VaR is [134]

$$VaR_\epsilon(\xi) \leq \sigma \sqrt{4/(9\epsilon) - 1} + \langle \omega, \xi \rangle = VaR_\epsilon^{VP}(\xi). \quad (10.2b)$$

The Cantelli bound is applicable for any probability distribution $\xi(\omega)$ and value $\epsilon \in [0, 1]$. The VP bound is sharper than the Cantelli bound, but is only valid when ξ is unimodal and $\epsilon \leq 1/6$.

10.2.2 Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space with time-indexed filtration \mathcal{F}_t , $X \subset \mathbb{R}^n$ be a compact set, and w be n -dimensional Wiener process. An Itô SDE with a drift function f and diffusion function g is [145]

$$dx = f(t, x)dt + g(t, x)dw. \quad (10.3)$$

In this chapter, trajectories will start from an initial set $X_0 \subseteq X$ and will remain within X in times $t \in [0, T]$ by virtue of stopping at the boundary ∂X . Define τ_X as a stopping time (random variable) corresponding to the time at which the process (10.3) starting from X_0 first touches the boundary ∂X for the first time. A process of (10.3) starting from an initial condition $x(0) \in X_0$ in times $t \in [0, T]$ is

$$x(t) = x(0) + \int_{t=0}^{\tau_X \wedge T} f(t, x)dt + \int_{t=0}^{\tau_X \wedge T} g(t, x)dw. \quad (10.4)$$

Solutions of (10.4) are unique if there exists finite constants $C, D > 0$ such that for all $(t, x, x') \in [0, T] \times X^2$, the following Lipschitz and Growth conditions hold [146]:

$$\begin{aligned} D\|x - x'\|_2 &\geq \|f(t, x) - f(t, x')\|_2 + \|g(t, x) - g(t, x')\|_2 \\ C(1 + \|x\|_2) &\geq \|f(t, x)\|_2 + \|g(t, x)\|_2. \end{aligned} \quad (10.5)$$

CHAPTER 10. VALUE-AT-RISK PEAK ESTIMATION

The Lipschitz and Growth conditions will hold if (f, g) are locally Lipschitz and the set X is compact. Distributions of the densities of (10.4) may be computed by solving a Fokker-Planck equation with absorbing boundary conditions on ∂X [147, 148].

The generator \mathcal{L} associated with the SDE is a linear operator that satisfies $\forall v(t, x) \in C^2([0, T] \times X)$ is

$$\mathcal{L}v(t, x) = \partial_t v + f(t, x) \cdot \nabla_x v + \frac{1}{2}g(t, x)^T (\nabla_{xx}^2 v) g(t, x). \quad (10.6)$$

The $\nabla_{xx}^2 v$ term arises from the Itô Lemma. Let τ be a stopping time adapted to the filtration, defined by $\tau = \tau_X \wedge T$. The occupation measure $\mu \in \mathcal{M}_+([0, T] \times X)$ corresponding to the stopping time τ , initial distribution $\mu_0 \in \mathcal{M}_+(X_0)$, and dynamics (10.3) is $\forall A \subseteq [0, T], B \subseteq X$ is

$$\mu(A \times B) = \int_{X_0} \int_{t=0}^{\tau} I_{A \times B}((t, x(t | x_0))) dt d\mu_0(x_0). \quad (10.7)$$

The initial measure $\mu_0 \in \mathcal{M}_+(X_0)$, the occupation measure μ from (10.7), and the terminal measure $\mu_\tau \in \mathcal{M}_+([0, T] \times X)$ defined by following (10.3) (from μ_0 until the stopping time τ) are all related by Dynkin's formula [149],

$$\langle v, \mu_\tau \rangle = \langle v(0, x), \mu_0(x) \rangle + \langle \mathcal{L}v, \mu \rangle \quad \forall v \in C^2. \quad (10.8)$$

Dynkin's formula is an SDE generalization of the Liouville equation for ODEs (2.8). Equation (10.8) may be equivalently written in weak form as

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu. \quad (10.9)$$

An expectation-maximizing optimal stopping problem for the SDE in (10.3) with a reward function of $p(x)$ in the region $[0, T] \times X$, when starting at the initial condition $x(0) \sim \mu_0 \in \mathcal{M}_+(X_0)$, is $P^* = \sup \mathbb{E}_{\mu_0}[p(x(\tau))]$. This expectation-maximization problem has the same expression as (3.2), but with a generator (10.6):

$$p^* = \sup \langle p, \mu_\tau \rangle \quad (10.10a)$$

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \quad (10.10b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (10.10c)$$

$$\mu, \mu_\tau \in \mathcal{M}_+([0, T] \times X) \quad (10.10d)$$

$$\mu_0 \in \mathcal{M}_+(X_0). \quad (10.10e)$$

Any μ that is part of a feasible solution (μ, μ_0, μ_τ) for (3.2b)-(3.2e) will be referred to as a *relaxed* occupation measure. Program (3.2) satisfies $p^* \geq P^*$, and tightness ($p^* = P^*$) is achieved under the assumptions of Lipschitz continuity and Growth (10.5), compactness of $[0, T] \times X$, and continuity of $p(x)$.

10.3 Peak Value-at-Risk Estimation for Stochastic Systems

This section will present the chance-peak problem statement, and will also derive the infinite-dimensional SOCP to upper bound the chance-peak quantile statistic.

10.3.1 Problem Statement

Let $\epsilon \in [0, 1]$ be a value for the quantile statistic, X be a compact set, $X_0 \subseteq X$ be a set of initial conditions, and (10.4) be the solution to an SDE evolving from $x(0) \in X_0$ that remains within X until it stops. For a given initial probability distribution $\mu_0 \in \mathcal{M}_+(X_0)$, and for all $t \in [0, T]$, let $x(t)$ be the stochastic process of (10.4) at time t , and let $\mu_t \in \mathcal{M}_+(X)$ be its probability distribution (with $x(t)$ stopping at ∂X).

10.3.1.1 Assumptions

The following assumptions will be posed throughout this chapter,

- A1 The set $[0, T] \times X$ is compact and $X_0 \subseteq X$.
- A2 The functions (f, g) satisfy (10.5).
- A3 The state function $p(x)$ is continuous on X .
- A4 The initial measure $\mu_0 \in \mathcal{M}_+(X_0)$ is a given probability distribution ($\langle 1, \mu_0 \rangle = 1$).

10.3.1.2 VaR Problem

Problem 10.3.1. *The chance-peak problem to find the ϵ -VaR of $p(x)$ is*

$$P^* = \sup_{t^* \in [0, T]} VaR_\epsilon(p \# \mu_{t^*}) \quad (10.11a)$$

$$dx = f(t, x)dt + g(t, x)dw \quad (10.11b)$$

from $t = 0$ until a stopping time of $\tau_X \wedge t^$* (10.11c)

$$x(0) \sim \mu_0. \quad (10.11d)$$

The pushforward $p \# \mu_{t^*}$ from (10.11a) is the univariate probability distribution of $p(x)$ at the state distribution $x \sim \mu_{t^*}$.

10.3.1.3 Tail-Bound Upper Bound

Let r be the constant factor multiplying σ in (10.2) such that

$$r^{cant} = \sqrt{1/(\epsilon) - 1} \quad r^{VP} = \sqrt{4/(9\epsilon) - 1}. \quad (10.12)$$

It is further assumed that the VP-bound will only be used if its conditions are satisfied ($\epsilon \leq 1/6$, unimodal). The distribution of $p(x)$ with respect to the state distribution μ_{t^*} is univariate, for which the relation in (10.1) and the constants in (10.2) can be used to upper-bound on Problem 10.3.1. We will use the notation $\langle p^2, \mu_{t^*} \rangle$ to refer to $\langle p(x)^2, \mu_{t^*}(x) \rangle$.

Problem 10.3.2. *The tail-bound program that upper-bounds the chance-peak (10.11) with constant r is*

$$P_r^* = \sup_{t^* \in [0, T]} r \sqrt{\langle p^2, \mu_{t^*} \rangle - \langle p, \mu_{t^*} \rangle^2} + \langle p, \mu_{t^*} \rangle \quad (10.13a)$$

$$dx = f(t, x)dt + g(t, x)dw \quad (10.13b)$$

from $t = 0$ until a stopping time of $\tau_X \wedge t^$* (10.13c)

$$x(0) \sim \mu_0. \quad (10.13d)$$

10.3.2 Nonlinear Measure Program

Problem 10.3.2 can be upper-bounded by an infinite-dimensional convex program in a given initial probability distribution μ_0 , terminal measure μ_τ , and relaxed occupation measure μ ,

using the generator \mathcal{L} in (6.13) as

$$p_r^* = \sup r \sqrt{\langle p^2, \mu_\tau \rangle - \langle p, \mu_\tau \rangle^2} + \langle p, \mu_\tau \rangle \quad (10.14a)$$

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \quad (10.14b)$$

$$\mu_\tau, \mu \in \mathcal{M}_+([0, T] \times X). \quad (10.14c)$$

Theorem 10.3.3. *Program 10.14 is an upper bound on (10.13) with $p_r^* \geq P_r^*$ under A1-A4.*

Proof. Let t^* be a stopping time in $[0, T]$, and let $x_0 \in X_0$ be an initial condition. Measures (μ_0, μ, μ_τ) that satisfy (10.14b) may be constructed from this (t^*, x_0) by μ_{t^*} as the state distribution of (10.4) at time t^* given μ_0 , and μ as the occupation measure in (10.7) associated to this SDE trajectory with distribution μ_0 . Because the feasible set to constraint (10.14b) contains measures induced by all possible provided SDE trajectories starting from μ_0 , it holds that $p_r^* \geq P_r^*$. \square

Remark 10.3.1. *The initial distribution $\mu_0 \in \mathcal{M}_+(X_0)$ may be optimized to find a supremal p_r^* over all probability distributions in X_0 variable by adding the constraint $\langle \mu_0, 1 \rangle = 1$ to (10.14).*

10.3.3 Measure Second-Order Cone Program

The nonlinear measure program (10.14) may be recast as an infinite-dimensional SOCP.

Lemma 10.3.4. *Let $J_r(a, b) = r\sqrt{b - a^2} + a$ be the objective (10.14a) with $a = \langle p(x), \mu_\tau \rangle$ and $b = \langle p(x)^2, \mu_\tau \rangle$. For any convex set $C \in \mathbb{R} \times \mathbb{R}_+$ with $(a, b) \in C$, the following pair of programs have the same optimal value (in which $Q^3 = \{(s, \kappa) \in \mathbb{R}^3 \times \mathbb{R}_+ \mid \|s\|_2 \leq \kappa\}$ is an SOC cone):*

$$\sup_{(a,b) \in C} a + r\sqrt{b - a^2} \quad (10.15)$$

$$\sup_{(a,b) \in C, z \in \mathbb{R}} a + rz : ([1 - b, 2z, 2a], 1 + b) \in Q^3. \quad (10.16)$$

Proof. The new variable z is introduced under the constraint $\sqrt{b - a^2} \geq z$, implying that $z^2 + a^2 \leq b$. The SOCP equivalence follows from the power-representation of $\sqrt{b - a^2}$ from [89, 150], with the steps of

$$([1 - b, 2z, 2a], 1 + b) \in Q^3 \quad (10.17a)$$

$$(1 - b)^2 + 4(z^2 + a^2) \leq (1 + b)^2 \quad (10.17b)$$

$$(1 + b^2) - 2b + 4(z^2 + a^2) \leq (1 + b^2) + 2b \quad (10.17c)$$

$$4(z^2 + a^2) \leq 4b. \quad (10.17d)$$

\square

Theorem 10.3.5. *An infinite-dimensional SOCP with the same optimal value and set of feasible solutions as (10.14) given μ_0 is*

$$p_r^* = \sup \quad rz + \langle p, \mu_\tau \rangle \quad (10.18a)$$

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \quad (10.18b)$$

$$u = [1 - \langle p^2, \mu_\tau \rangle, 2z, 2\langle p, \mu_\tau \rangle] \quad (10.18c)$$

$$(u, 1 + \langle p^2, \mu_\tau \rangle) \in Q^3 \quad (10.18d)$$

$$\mu, \mu_\tau \in \mathcal{M}_+([0, T] \times X), z \in \mathbb{R}, u \in \mathbb{R}^3. \quad (10.18e)$$

Proof. This results from an application of Lemma 10.3.4 to the objective term (10.14a). The optimization variables are now (μ_τ, μ, z, u) . \square

Corollary 3. *Program (10.18) is convex.*

Proof. The objective (10.18a) is affine in (z, μ_τ) . Constraints (10.18b)-(10.18e) are convex (affine for (10.18b) and SOC for (10.18d)), ensuring convexity of (10.18). \square

Remark 10.3.2. *Problem (10.18) has an infinite-dimensional affine constraint in (10.18b) and a finite-dimensional SOC constraint in (10.18d)*

10.3.4 Dual Second-Order Cone Program

The Lagrangian dual of (10.18) is a program with infinite-dimensional linear constraints and a finite-dimensional SOC constraint. This dual involves a function $v(t, x) \in C^2([0, T] \times X)$ and a constant $y \in \mathbb{R}^3$ as variables.

We will use the following expression of (10.18) with an explicitly written SOC variable q linked with linear constraints to the measures (μ_0, μ_τ, μ) .

Lemma 10.3.6. *The following program has the same optimal value as (10.18):*

$$p_r^* = \sup (r/2)q_2 + \langle p, \mu_\tau \rangle \quad (10.19a)$$

$$q_1 + \langle p^2, \mu_\tau \rangle = 1 \quad (10.19b)$$

$$q_3 - 2\langle p, \mu_\tau \rangle = 0 \quad (10.19c)$$

$$q_1 + q_4 = 2 \quad (10.19d)$$

$$\mu_\tau - \mathcal{L}^\dagger \mu = \delta_0 \otimes \mu_0 \quad (10.19e)$$

$$q = ([q_1, q_2, q_3], q_4) \in Q^3 \quad (10.19f)$$

$$\mu, \mu_\tau \in \mathcal{M}_+([0, T] \times X). \quad (10.19g)$$

Proof. This formulation is obtained from (10.18) through the change of variable $q = (u, 1 + \langle p^2, \mu_\tau \rangle)$, the replacement of z with $q_2/2$ using the second coordinate of constraint (10.18c), adding the first coordinate of (10.18c) and the last coordinate of (10.18d) to derive (10.19d).

□

Theorem 10.3.7. *The dual program of (10.18) with weak duality $d_r^* \geq p_r^*$ under Assumptions A1-A3 is*

$$d_r^* = \inf y_1 + 2y_3 + \int_{X_0} v(0, x_0) d\mu_0(x_0) \quad (10.20a)$$

$$\mathcal{L}v(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (10.20b)$$

$$v(t, x) + y_1 p^2(x) - 2y_2 p(x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (10.20c)$$

$$([y_1 + y_3, -(r/2), y_2], y_3) \in Q^3 \quad (10.20d)$$

$$y \in \mathbb{R}^3, v \in C^2([0, T] \times X). \quad (10.20e)$$

Strong duality with $d_r^ = p_r^*$ holds under Assumptions A1-A4.*

Proof. **Dual formulation:** this formulation is obtained by applying the standard Lagrangian duality method to (10.19). v is the Lagrange multiplier corresponding to constraint (10.19e), and y is the Lagrange multiplier corresponding to constraints (10.19b)-(10.19d). Conversely, μ is the Lagrange multiplier corresponding to constraint (10.20b), μ_τ is the Lagrange multiplier corresponding to constraint (10.20c), and q is the Lagrange multiplier corresponding to (10.20d). The cost in (10.20a) corresponds to the right-hand sides of constraints (10.19b)-(10.19e), while the right-hand side of (10.20c) and the second coordinate $-(r/2)$ in (10.20d) correspond to the cost in (10.19a).

Strong duality: see Appendix A.8. □

This strong duality property is an important feature of the infinite-dimensional problem at hand: it means that one may equivalently solve moment relaxations of (10.19) and sums-of-squares relaxations of (10.20).

10.4 Finite Moment Program

This section will upper-bound (10.18) utilizing a converging hierarchy of SDPs in increasing degree and size.

10.4.1 Moment Program

The following assumptions are required to utilize the moment-SOS hierarchy in approximating (10.18):

A5 The sets X_0 and X are Archimedean BSA sets.

A6 The functions $f(t, x), g(t, x)$ are polynomial vectors and $p(x)$ is a polynomial scalar.

Let $(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^\tau)$ be moment sequences corresponding to the measures (μ_0, μ, μ_τ) respectively. For each monomial $x^\alpha t^\beta$ with $\alpha \in \mathbb{N}^n, \beta \in \mathbb{N}$, define the operator $\text{Dyn}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}, \mathbf{m}^\tau)$ as the linear relation induced by Dynkin's formula (10.9),

$$\langle x^\alpha, \mu_0 \rangle \delta_{\beta 0} + \langle \mathcal{L}(x^\alpha t^\beta), \mu \rangle - \langle x^\alpha t^\beta, \mu_\tau \rangle = 0. \quad (10.21)$$

Define the dynamics degree \tilde{d} as

$$\tilde{d} = d + \max(\lfloor \deg f/2 \rfloor, \deg g - 1). \quad (10.22)$$

Problem 10.4.1. For $d \geq \deg(p)$, the order- d moment problem that upper-bounds problem (10.18), with variables $(\tau,)$ given μ_0 is

$$p_{r,d}^* = \max_{z \in \mathbb{R}} \quad rz + L_{\mathbf{m}^\tau}(p(x)) \quad (10.23a)$$

$$\text{Dyn}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}^\tau, \mathbf{m}) = 0 \quad \forall (\alpha, \beta) \in \mathbb{N}_{\leq 2d}^{n+1} \quad (10.23b)$$

$$s = [1 - L_{\mathbf{m}^\tau}(p(x))^2, 2z, 2L_{\mathbf{m}^\tau}(p(x))] \quad (10.23c)$$

$$(s, 1 + L_{\mathbf{m}^\tau}(p(x))^2) \in Q^3 \quad (10.23d)$$

$$\mathbb{M}_d(([0, T] \times X)\mathbf{m}^\tau) \succeq 0 \quad (10.23e)$$

$$\mathbb{M}_{\tilde{d}}(([0, T] \times X)\mathbf{m}) \succeq 0. \quad (10.23f)$$

Constraint (10.23b) is a finite-dimensional truncation of the infinite-dimensional (10.18b).

The following boundedness result is required to ensure convergence:

Lemma 10.4.2. *All of (μ, μ_τ, z) are bounded in (10.18) under A1-A3.*

Proof. A sufficient condition for a measure to be bounded (all moments are bounded) is that it has finite mass and is supported on a compact set. Compactness of $[0, T] \times X$ holds by A1. Assumption A1 imposes that $\langle 1, \mu_0 \rangle = 1$. By substituting $v(t, x) = 1$ to (10.8) in (10.18b), it holds that $\langle 1, \mu_\tau \rangle = \langle 1, \mu_0 \rangle = 1$. Performing the same step with $v(t, x) = t$ yields $T \geq \langle t, \mu_\tau \rangle = \langle 1, \mu \rangle$. It therefore holds that $\langle p, \mu_\tau \rangle$ and $\langle p^2, \mu_\tau \rangle$ are bounded, and $\langle p^2, \mu_\tau \rangle$ is bounded from below by 0. The SOC constraint (10.18d) ensures that z is finite, demonstrating that all variables are bounded. \square

Remark 10.4.1. *The relation $p_d^* \geq p_r^* \geq P_r^*$ will still hold when $[0, T] \times X$ is noncompact (violating A1 and A5), but it may no longer occur that $\lim_{d \rightarrow \infty} p_{r,d}^* = p_r^*$ (the conditions Lemma 10.4.3 will no longer apply).*

Theorem 10.4.3. *Under assumptions A1-A6, (10.23) inherits the strong duality property of its infinite-dimensional counterpart (10.18), and its optima will converge to (10.18) i.e. $\lim_{d \rightarrow \infty} p_{r,d}^* = p_r^*$.*

Proof. Strong duality is proved almost identically in the finite dimensional setting as in the infinite dimensional setting of Theorem 10.3.7 by using the same arguments as in the proof of [23, Proposition 6].

Convergence is a direct consequence of [23, Corollary 8] (when extending to the case with finite-dimensional SOC variables) through Lemma 10.4.2. \square

Remark 10.4.2. *The relation $p_d^* \geq p_r^* \geq P_r^*$ will still hold when $[0, T] \times X$ is noncompact (violating A1 and A5), but it may no longer occur that $\lim_{d \rightarrow \infty} p_{r,d}^* = p_r^*$ (the conditions Lemma 10.4.3 will no longer apply).*

10.4.2 Computational Complexity

At order- d , the size of the moment matrices corresponding to the measures is described in Table 10.1.

The LMI in (10.23) must be converted to SDP-standard form by introducing equality constraints between the entries of the moment matrices in order to utilize symmetric-cone Interior Point Methods (e.g., Mosek [49]). The per-iteration complexity of an SDP involving a single moment matrix of size $\binom{n+d}{d}$ scales as n^{6d} [140]. The scaling of an SDP with multiple moment and localizing

Table 10.1: Size of Moment Matrices in LMI (10.23)

$$\begin{array}{ll} \text{Matrix:} & \mathbb{M}_{\tilde{d}}(\mathbf{m}) \quad \mathbb{M}_d(\mathbf{m}^\tau) \\ \text{Size:} & \binom{n+1+\tilde{d}}{\tilde{d}} \quad \binom{n+1+d}{d} \end{array}$$

matrices generally depends on the maximal size of any PSD matrix. By Table 10.1, this size is $\binom{n+1+\tilde{d}}{\tilde{d}}$ with a scaling impact of $(n+1)^{6\tilde{d}}$. The complexity of using this chance-peak routine increases in a jointly polynomial manner with the degree \tilde{d} and the number of states n .

10.5 Extensions

This section outlines extensions to the developed chance-peak framework.

10.5.1 Exit-Time Statistics

This extension builds upon the work of [151] in estimating expectations of functions upon first exit of the region X . The exit time distribution of the SDE in (10.4) is the class of trajectories that stop according to the stopping time τ_X . The program (3.2) is modified in [151] by adding a support constraint $\mu_\tau \in \mathcal{M}_+(\partial X)$. Averaged statistics of the exit time distribution (expectations of $p_\# \mu_\tau$) may be computed by substituting different choices of p into the ∂X -adjusted (3.2), and this expectation is bounded from above and below with no relaxation gap in [151]. An example of this kind of statistic is finding the mean (stopping) time at which the trajectory exits by estimating $\langle t, \mu_\tau \rangle$.

We can adapt this methodology to bound $1 - \epsilon$ quantile values of the exit time distribution by restricting $\mu_\tau \in \mathcal{M}_+(\partial X)$ constraint in Program (10.18) and (10.18e) by

$$p_r^* = \sup \quad rz + \langle p, \mu_\tau \rangle \tag{10.24a}$$

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \tag{10.24b}$$

$$u = [1 - \langle p^2, \mu_\tau \rangle, 2z, 2\langle p, \mu_\tau \rangle] \tag{10.24c}$$

$$(u, 1 + \langle p^2, \mu_\tau \rangle) \in Q^3 \tag{10.24d}$$

$$\mu \in \mathcal{M}_+([0, T] \times X), z \in \mathbb{R}, u \in \mathbb{R}^3. \tag{10.24e}$$

$$\mu_\tau \in \mathcal{M}_+([0, T] \times \partial X). \tag{10.24f}$$

The optimal value (10.24) is a (Cantelli or VP) upper-bound on the exit time quantity $VaR_\epsilon(p_\# \mu_\tau)$.

10.5.2 Distance Estimation

The chance-peak methodology developed in this chapter can be applied towards bounding (probabilistically) the distance of closest approach to an unsafe set (using the framework from Chapter 5). Let $X_u \subset X$ be an unsafe set, and let $c(x, y)$ be a metric in X . The point-set distance function with respect to X_u is $c(x; X_u) = \inf_{y \in X_u} c(x, y)$. The output of a chance-distance program is an infimal value c^* such that there exists some time in which the probability of traveling closer than c^* to X_u is at least ϵ .

The chance-distance C^* is the negative of the bound P^* obtained from solving (10.11) with objective $p(x) = -c(x; X_u)$. Because the objective $c(x; X_u)$ is not generally polynomial (even when $c(x, y)$ is polynomial), the LMI (10.23) cannot directly be posed in terms of $c(x; X_u)$. One method to maintain a polynomial structure is to add time-constant states $dy = \mathbf{0}dt + \mathbf{0}dw$ to dynamics (10.3) in x and form the state support set $(x, y) \in X \times X_u$. When X_u is full-dimensional inside $X \subset \mathbb{R}^n$, the occupation measure $\mu(t, x, y) \in \mathcal{M}_+([0, T] \times X \times Y)$ will have a moment matrix of size $\binom{1+2n+d}{d}$ at each fixed degree d .

This size can be reduced by decomposing the peak measure $\hat{\mu}_\tau \in \mathcal{M}_+([0, T] \times X \times Y)$ into a joint measure $\eta(x, y) \in \mathcal{M}_+(X \times Y)$ and a peak measure $\mu_\tau(t, x) \in \mathcal{M}_+([0, T] \times X)$ that are equal in the x marginal. The chance-distance SOCP under this decomposition is

$$p_r^* = \sup_{z \in \mathbb{R}} \quad rz + \langle -c(x, y), \eta(x, y) \rangle \quad (10.25a)$$

$$\langle v, \mu_\tau \rangle = \langle v(0, x_0), \mu_0(x_0) \rangle + \langle \mathcal{L}v, \mu \rangle \quad \forall v \in C^2([0, T] \times X) \quad (10.25b)$$

$$\langle \theta(x), \mu_\tau(t, x) \rangle - \langle \theta(x), \eta(x, y) \rangle = 0 \quad \forall \theta \in C(X) \quad (10.25c)$$

$$s = [1 - \langle c^2, \eta \rangle, 2z, 2\langle -c, \eta \rangle] \quad (10.25d)$$

$$(s, 1 + \langle c^2, \eta \rangle) \in Q^3 \quad (10.25e)$$

$$\eta \in \mathcal{M}_+(X \times X_u) \quad (10.25f)$$

$$\mu, \mu_\tau \in \mathcal{M}_+([0, T] \times X). \quad (10.25g)$$

Constraint (10.25c) enforces equality in the x -marginals between μ_τ and η . The Moment matrices of η and μ respectively in the LMI program derived from (10.25) have sizes $\binom{2n+d}{d}$ and $\binom{n+1+\tilde{d}}{\tilde{d}}$. Unfortunately, the squaring operation $\langle c^2, \eta \rangle$ causes mixed multiplications in variables even when c is additively separable as $c(x, y) = \sum_{i=1}^n c_i(x_i, y_i)$, thus forbidding the application of correlative sparsity ([61] and Section 5.6) to reduce the complexity of LMIs from (10.25).

10.5.3 Switching

The chance-peak scheme may also be applied to switched stochastic systems. The methods outlined in this section are an extension of the ODE approach from Section 4.2.3, and are similar to duals of constraints found in [141]. Assume that there are $N_s \in \mathbb{N}$ subsystems indexed by $\ell = 1..N_s$. Each subsystem has individual dynamics

$$dx = f_\ell(t, x)dt + g_\ell(t, x), \quad (10.26)$$

for each switching index $\ell = 1..N_s$. A switched SDE trajectory is a distribution $x(t)$ and a switching function $S : [0, T] \rightarrow (1..N_s)$ under the constraint that $x(t)$ satisfies (10.26) whenever $S(t) = \ell$ (the ℓ -th subsystem is active). A specific trajectory of a switched SDE starting from an initial point $x_0 \in X$ will be expressed as $x(t | x_0, S)$. No dwell time constraints are imposed on the switching sequence S ; instead, switching can occur arbitrarily quickly in time.

Generators \mathcal{L}_ℓ may be defined for each subsystem in (10.26) according to $\forall v \in C^2([0, T] \times X :$

$$\mathcal{L}_\ell v(t, x) = \partial_t v + f_\ell \cdot \nabla_x v + g_\ell^T(\nabla_{xx}^2 v)g_\ell/2 \quad \forall \ell = 1..N_s. \quad (10.27)$$

Let $\mu \in \mathcal{M}_+([0, T] \times X)$ be the total occupation measure of the switched SDE trajectory $x(t | x_0, S)$. The total occupation measure may be split into disjoint subsystem occupation measures $\forall \ell : \mu_\ell \in \mathcal{M}_+([0, T] \times X)$ under the relation $\sum_{\ell=1}^{N_s} \mu_\ell = \mu$. The mass of a subsystem's occupation measure $\langle 1, \mu_\ell \rangle$ is the total amount of time that the trajectory $x(t | x_0, S)$ spends in subsystem $S(t) = \ell$.

Dynkin's equation (10.9) for switching-type uncertainty is

$$\mu_\tau = \delta_0 \otimes \mu_0 + \sum_{\ell=1}^{N_s} \mathcal{L}_\ell^\dagger \mu_\ell. \quad (10.28)$$

The chance-peak problem in (10.18) modified for switching uncertainty is

$$p_r^* = \sup \quad rz + \langle p, \mu_\tau \rangle \quad (10.29a)$$

$$\mu_\tau = \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \quad (10.29b)$$

$$u = [1 - \langle p^2, \mu_\tau \rangle, 2z, 2\langle p, \mu_\tau \rangle] \quad (10.29c)$$

$$(u, 1 + \langle p^2, \mu_\tau \rangle) \in Q^3 \quad (10.29d)$$

$$\mu_\tau \in \mathcal{M}_+([0, T] \times X), z \in \mathbb{R}, u \in \mathbb{R}^3. \quad (10.29e)$$

$$\mu_\ell \in \mathcal{M}_+([0, T] \times \partial X) \quad \forall \ell = 1..N_s. \quad (10.29f)$$

10.6 Numerical Examples

MATLAB (2022a) code to replicate experiments is available at https://github.com/Jarmill/chance_peak. Dependencies include Mosek [49] and YALMIP [68].

10.6.1 Two States

Example 1 of [152] is the following two-dimensional cubic polynomial SDE

$$dx = \begin{bmatrix} x_2 \\ -x_1 - x_2 - \frac{1}{2}x_1^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} dw. \quad (10.30)$$

This example performs chance-peak maximization of $p(x) = -x_2$ starting at the point (Dirac-delta initial measure μ_0) $X_0 = [1, 1]$ with $X = [-1, 2] \times [-1, 1.5]$ and $T = 5$. Trajectories of (10.30) are displayed in cyan in Figure 10.1 starting from the black-circle X_0 , and four of these trajectories are marked in non-cyan colors. The $\epsilon = 0.5$ row of Table 10.2 displays the bounds on the mean distribution as solved through finite-degree SDP truncations of (3.2). The bounds at $\epsilon = \{0.15, 0.1, 0.05\}$ are obtained through the VP expression in (10.2b) and solving the SDPs obtained from (10.23). The dotted and solid red lines in Figure 10.1 are the $\epsilon = 0.5$ and $\epsilon = 0.15$ bounds respectively at order 5.

Table 10.2: Chance-Peak estimation of the Stochastic Flow System (10.30) to maximize $p(x) = -x_2$

order	1	2	3	4	5	6
$\epsilon = 0.5$	1.5000	0.8817	0.8773	0.8747	0.8745	0.8744
$\epsilon = 0.15$	1.2910	1.2210	1.1817	1.1689	1.1657	1.1642
$\epsilon = 0.1$	1.5811	1.5009	1.4520	1.4361	1.4323	1.4303
$\epsilon = 0.05$	2.2361	2.1306	2.0613	2.0387	2.0332	2.0305

Table 10.3: Solver time (seconds) to compute Table 10.2

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.643	0.417	0.431	1.963	1.659	3.641
$\epsilon = 0.15$	0.272	0.216	0.325	1.592	4.178	7.464
$\epsilon = 0.1$	0.264	0.213	0.316	1.651	1.339	4.225
$\epsilon = 0.05$	0.184	0.222	0.366	0.936	2.446	5.298

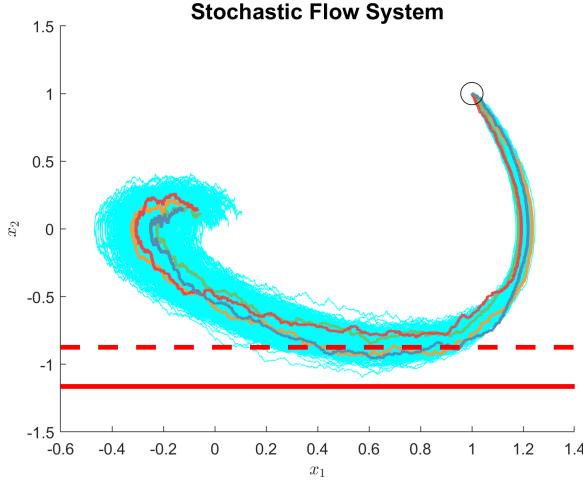


Figure 10.1: Trajectories of (10.30) with $\epsilon = \{0.5, 0.15\}$ bounds

10.6.2 Three States

An SDE modification of the Twist system (5.34) is

$$dx = \begin{bmatrix} -2.5x_1 + x_2 - 0.5x_3 + 2x_1^3 + 2x_3^3 \\ -x_1 + 1.5x_2 + 0.5x_3 - 2x_2^3 - 2x_3^3 \\ 1.5x_1 + 2.5x_2 - 2x_3 - 2x_1^3 - 2x_2^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ 0.1 \end{bmatrix} dw. \quad (10.31)$$

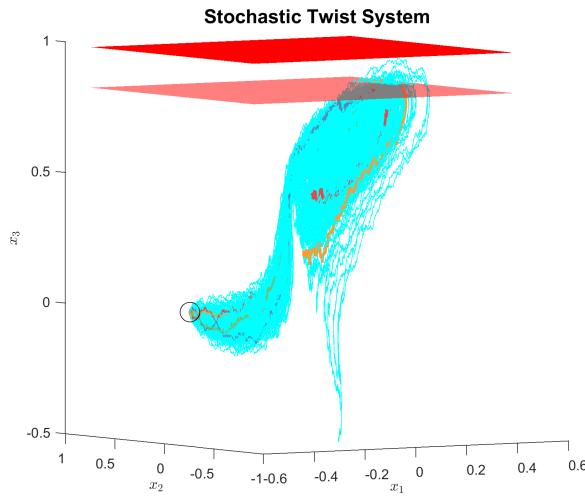
This second example performs chance-peak maximization of $p(x) = x_3$ starting at the point $X_0 = [0.5, 0, 0]$ with $X = [-0.6, 0.6] \times [-1, 1] \times [-1, 1.5]$ and $T = 5$. VP bounds from solving the SDEs from (3.2) and (10.23) are recorded in Table 10.4 in the same manner as in Table 10.2. Figure 10.2 plots trajectories and bounds of (10.31) starting from the black-circle X_0 point, with four of these trajectories visibly distinguished. The solid red plane in Figure 10.2 is the $\epsilon = 0.15$ bound on x_3 at order 6, and the translucent red plane is the $\epsilon = 0.5$ bound on x_3 (also at order 6).

Table 10.4: Chance-Peak estimation of the Stochastic Twist System (10.31) to maximize $p(x) = x_3$

order	1	2	3	4	5	6
$\epsilon = 0.5$	1.4682	0.9100	0.8312	0.8231	0.8211	0.820254
$\epsilon = 0.15$	1.2910	1.1729	1.0813	1.0171	0.9891	0.975507
$\epsilon = 0.1$	1.5811	1.4358	1.3180	1.2217	1.1859	1.169364
$\epsilon = 0.05$	2.2361	2.0288	1.8497	1.6866	1.6299	1.589447

Table 10.5: Solver time (seconds) to compute Table 10.4

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.734	0.526	1.880	5.213	28.269	81.736
$\epsilon = 0.15$	0.316	0.397	1.117	3.876	23.302	132.539
$\epsilon = 0.1$	0.144	0.210	1.080	6.313	27.829	116.733
$\epsilon = 0.05$	0.118	0.185	1.352	4.999	20.863	137.833


 Figure 10.2: Trajectories of (10.31) with $\epsilon = 0.5$ (transparent red) and $\epsilon = 0.15$ (solid red) bounds

10.6.3 Exit-Time

This example uses the setting of Example 7.4 of [151]. The dynamics are standard Brownian motion in $n = 3$ dimensions with $f = \mathbf{0}_3$ and $g = \mathbf{1}_3$. The initial condition is $X_0 = \mathbf{0}_3$ with a support set of $X = \{x \in \mathbb{R}^3 \mid \sum_{i=1}^3 x_i^4 \leq 1\}$. The considered boundary is $\partial X = \{x \in \mathbb{R}^3 \mid \sum_{i=1}^3 x_i^4 = 1\}$.

Chance-peak bounds for the first arrival time $\langle t, \mu_\tau \rangle$ are displayed in Table 10.6. The order-1 estimates are dual infeasible and are marked by ∞ .

Table 10.6: Chance-Peak Exit-Statistic estimation of Standard Brownian Motion System

order	1	2	3	4	5
$\epsilon = 0.5$	∞	0.0642	0.0642	0.0642	0.0642
$\epsilon = 0.15$	∞	0.1605	0.1088	0.1088	0.1088
$\epsilon = 0.1$	∞	0.2001	0.1233	0.1233	0.1233
$\epsilon = 0.05$	∞	0.2845	0.1537	0.1537	0.1537

10.6.4 Distance Estimation

This example will involve distance estimation of the following SDE based on [36]

$$dx = \begin{bmatrix} x_2 \\ -x_1 - x_2 + \frac{1}{3}x_1^3 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} dw. \quad (10.32)$$

This L_2 chance-distance task takes place at a time horizon of $T = 5$ with sets $X_0 = [1; 1]$, $X = [-1.1, 1.75] \times [-1.5, 1.5]$, and $X_u = \{y \in \mathbb{R}^2 \mid 0.5^2 \geq (y_1 + 0.5)^2 + (y_2 + 0.75)^2\}$. Distance estimation was accomplished by maximizing VaRs of the function $-\|x - y\|_2^2$ in (10.25).

System trajectories of (10.32) are displayed in Figure 10.3, in which the unsafe half-circle set X_u is drawn in solid red. Squared distance lower bounds from solving SDPs arising from moment programs of (10.25) are listed in Table 10.7. Negative distance lower-bounds are truncated to 0 in Table 10.7.

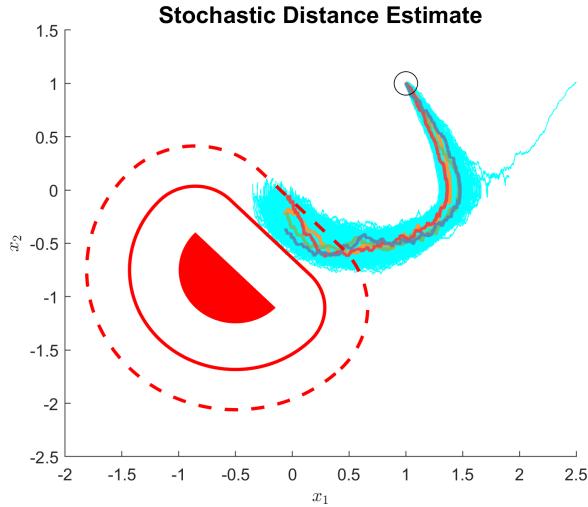


Figure 10.3: Trajectories of (10.32) with $\epsilon = \{0.5, 0.15\}$ bounds

Table 10.7: Chance-Peak squared distance lower bounds for System (10.32)

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.0328	0.3989	0.5095	0.6186	0.6559	0.7055
$\epsilon = 0.15$	0	0	0	0.0860	0.1872	0.3685
$\epsilon = 0.1$	0	0	0	0	0.0333	0.2508
$\epsilon = 0.05$	0	0	0	0	0	0

Table 10.8: Solver time (seconds) to compute Table 10.7

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.759	0.439	0.588	1.552	7.317	22.313
$\epsilon = 0.15$	0.306	0.325	0.671	1.380	6.091	26.481
$\epsilon = 0.1$	0.270	0.284	0.739	1.751	7.249	21.590
$\epsilon = 0.05$	0.291	0.282	0.578	1.326	6.199	26.908

10.6.5 Switching

We utilize a modification of Example C from [141] for this final example. The two subsystems involved are,

$$dx = \begin{bmatrix} -2.5x_1 - 2x_2 \\ -0.5x_1 - x_2 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.25x_2 \end{bmatrix} dw \quad (10.33a)$$

$$dx = \begin{bmatrix} -x_1 - 2x_2 \\ 2.5x_1 - x_2 \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0.25x_2 \end{bmatrix} dw. \quad (10.33b)$$

Switched SDE trajectories start from an initial condition of $X_0 = (0, 1)$ and are tracked in the state set $X = [-2, 2]^2$ with a time horizon of $T = 5$. The chance-peak problem is solved to find bounds on $p(x) = -x_2$.

Figure (10.4) plots switched SDE trajectories along with $\epsilon = \{0.5, 0.15\}$ bounds (at order-6). Table 10.9 lists these discovered bounds.

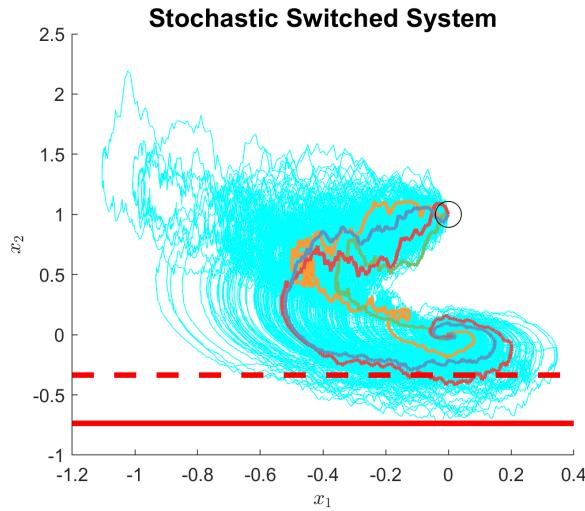

 Figure 10.4: Trajectories of (10.33) with $\epsilon = \{0.5, 0.15\}$ bounds

Table 10.9: Chance-Peak upper-bounds for $p(x) = -x_2$ for the Switched System (10.33)

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.8491	0.4304	0.3823	0.3630	0.3488	0.3350
$\epsilon = 0.15$	1.1619	0.8214	0.7736	0.7539	0.7416	0.7361
$\epsilon = 0.1$	1.4301	1.0452	0.9903	0.9675	0.9534	0.9469
$\epsilon = 0.05$	2.0341	1.5460	1.4749	1.4449	1.4268	1.4189

Table 10.10: Solver time (seconds) to compute Table 10.9

order	1	2	3	4	5	6
$\epsilon = 0.5$	0.703	0.433	0.252	0.478	1.511	2.943
$\epsilon = 0.15$	0.148	0.131	0.185	0.522	1.088	2.463
$\epsilon = 0.1$	0.106	0.129	0.188	0.481	1.468	2.687
$\epsilon = 0.05$	0.111	0.128	0.184	0.669	1.695	2.963

10.7 Conclusion

This chapter considered the chance-peak problem, which involved finding upper bounds on the quantiles of state functions $p(x)$ achieved by SDE systems. The true $(1 - \epsilon)$ -quantile statistic P^* (10.11) is upper-bounded by the Cantelli/VP approximation P_r^* (10.13), which in turn is upper bounded by an infinite-dimensional SOCP p_r^* (10.18) and its moment-SOS finite-dimensional SDPs, yielding $p_{r,d}^*$ with $\lim_{d \rightarrow \infty} p_{r,d}^* = p_d^*$. Each of these upper-bounds contribute valuable information towards the analysis of SDEs.

Future work includes finding conditions under which the measure-based upper-bounding does not add conservatism (e.g., cases where $p_r^* = P_r^*$), and utilizing higher-moment tail-probability inequalities to obtain closer estimates to the VaR [153]. Another avenue involves developing stochastic optimal control strategies to minimize quantile statistics. Other aspects could include extension of Switching methods towards more general Lévy processes [154, 139].

Chapter 11

Peak Estimation for Time-Delay Systems

11.1 Introduction

This chapter presents an algorithm to upper bound extreme values of a state function attained along trajectories of a Delay Differential Equation (DDE) (time-delay system). The dynamics of a DDE depend on a history of the state, in contrast to an ODE in which the dynamics are a function only of the present values of state [155, 156, 157, 158]. This chapter will involve analysis of DDEs in a state space $X \subset \mathbb{R}^n$ over a time horizon $T < \infty$ with a single fixed discrete bounded delay $\tau \in (0, T)$.

Trajectory evolution of a DDE depends on an initial history $x_h : [-\tau, 0] \rightarrow X$ rather than simply an initial condition $x_0 \in X$ for a corresponding ODE. The evaluation at time t for a trajectory starting with a history x_h will be denoted as $x(t | x_h)$. A function class \mathcal{H} of histories may be defined, allowing for the definition of differential inclusions of DDEs. A peak estimation problem may be defined on a time-delay system to find the maximum value of a state function p along system trajectories given a class of initial histories \mathcal{H} as

$$P^* = \sup_{t^* \in [0, T], x_h(\cdot)} p(x(t^* | x_h)) \quad (11.1a)$$

$$\dot{x} = f(t, x(t), x(t - \tau)) \quad \forall t \in [0, T] \quad (11.1b)$$

$$x(t) = x_h(t) \quad \forall t \in [-\tau, 0] \quad (11.1c)$$

$$x_h(\cdot) \in \mathcal{H}. \quad (11.1d)$$

The variables in Problem (11.1) are the stopping time t^* and the initial history x_h . Problem (11.1) is a DDE version of the (generically nonconvex) ODE peak estimation program studied in

CHAPTER 11. PEAK ESTIMATION FOR TIME-DELAY SYSTEMS

[5, 6]. The peak estimation task in (11.1) is an instance of a DDE OCP with a free terminal time and a zero running (integrated) cost.

This work uses measure-theoretic methods in order to provide certifiable upper bounds on the peak value P^* from (11.1). The first application of measure-theoretic methods towards DDEs was in [159], in which the control input was relaxed into a Young Measure [45] (probability distribution at each point in time) [160]. This Young-Measure-based relaxed control yields the OCP optimal value in the case of a single discrete time delay under convexity, regularity, and compactness assumptions. However, the Young Measure control programs may result in a lower bound when there are two or more delays in the system dynamics (there exist Young-Measure solutions that do not correspond to OCP solutions) [161, 162]. Adding new measures and constraints allows for the construction of tight Young Measure OCP approximations at the cost of significantly more complicated programs [163].

Use of the moment-SOS hierarchy towards analysis of DDEs includes finding stability and safety certificates [164, 165, 166]. Prior work on using occupation measures for problems in time delays includes ODE-Partial Differential Equation (PDE) models in [167, 168], a Riesz-frame system in [169], and a gridded LP framework for optimal control given a single history x_h in [170]. Peak estimation has been conducted on specific time-delay systems such as the forced Liénard model [171] and compartmental epidemic models [11].

This chapter is organized as follows: Section 11.2 formalizes notation and summarizes concepts in measure theory, time-delay, occupation measures, and ODE peak estimation. Section 11.3 defines Measure-Valued (MV)-solutions for free-terminal-time DDE trajectories to create a primal-dual pair of LPs in order to upper-bound (11.1). Section 11.4 applies the Moment-SOS hierarchy towards generating SDPs to upper-bound the peak-estimation measure LP. Section 11.5 extends the DDEs peak framework by allowing for distance estimation, shaping constraints on histories, and uncertainty. Section 11.6 provides examples of DDE peak estimation. Section 11.7 concludes the chapter. Appendix A.9 extends the MV-solution framework towards continuous-time systems with proportional delays and discrete-time systems with long time delays. Appendix A.10 performs the proof of strong duality for the DDEs peak estimation LPs. Appendix A.11 finds and analyzes structural properties of DDE OCPs subvalue functionals. Appendix A.12 reduces conservatism of OCP approximations by performing spatio-temporal partitioning and applying double-integral subvalue functionals. Appendix A.13 introduces a more conservative but computationally simpler notion of MV-solution for DDEs.

This work appeared in [172] and is coauthored by Milan Korda, Victor Magron, and Mario Sznaier.

11.2 Time Delay Background

A single-variable function $g(t)$ is Piecewise Continuous (PC) over the domain $[a, b]$ if there exist $B \in \mathbb{N} \setminus \{0\}$ and a finite number of time-breaks $t_0 = a < t_1 < t_2 < \dots < t_B < b = t_{B+1}$ such that the function $g(t)$ is continuous in each interval $[t_k, t_{k+1})$ for $k = 0..B$. The class of PC functions from the time interval $[-\tau, 0]$ to X is $PC([-\tau, 0], X)$.

Given a PC state history $t \mapsto x_h(t)$, $t \in [-\tau, 0]$, a unique forward trajectory $x(t | x_h)$ of (11.1b) exists on $t \in [0, T]$ if the function $(t, x_0, x_1) \mapsto f(t, x_0, x_1)$ is locally Lipschitz in all variables.

Trajectories of time-delay systems with the form of (11.1b) with f locally Lipschitz satisfy a smoothing property as shown in Figure 11.1. The order of derivatives that are continuous will increase by 1 every τ time steps [158]. An example of such a time-delay system with increasing continuity is visualized in Figure 11.1 with system dynamics

$$x'(t) = -2x(t) - 2x(t - 1). \quad (11.2)$$

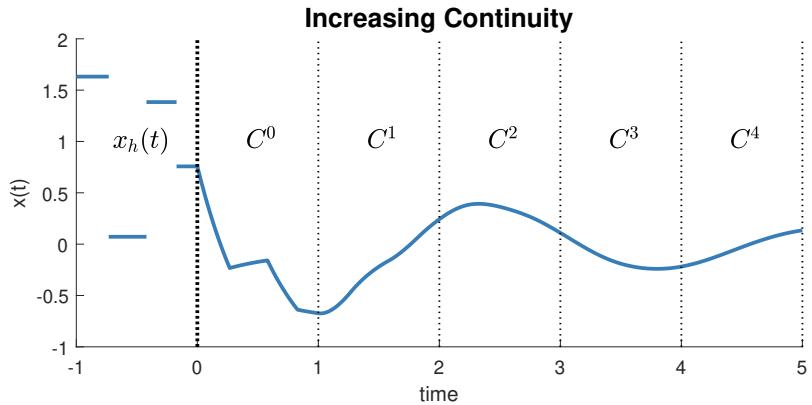


Figure 11.1: Continuity of (11.2) trajectories increases every $\tau = 1$ time step

Figure 11.2 plots multiple trajectories of (11.2) whose histories are lines passing through $x_h(0) = 1$, but whose evolution after time $t = 0$ is different.

The behavior of time-delay systems may change and bifurcate as the time delays change. A well-studied example of $\dot{x} = -x(t - \tau)$ is plotted in Figure (11.3) [158], in which the system is stable (to $x = 0$) for all bounded PC histories with $\tau \in [0, \pi/2)$, has bounded oscillations for some initial histories at $\tau = \pi/2$ (e.g., constant x_h in time), and is unstable (divergent oscillations to $\pm\infty$) for all similar histories with $\tau > \pi/2$ [158].

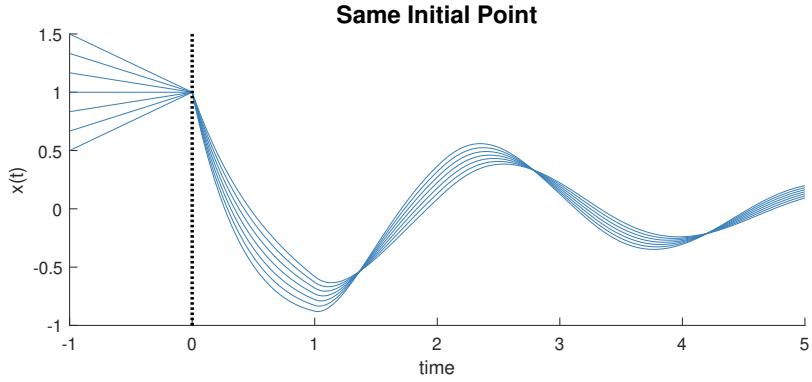


Figure 11.2: All histories pass through $x_h(0) = 1$

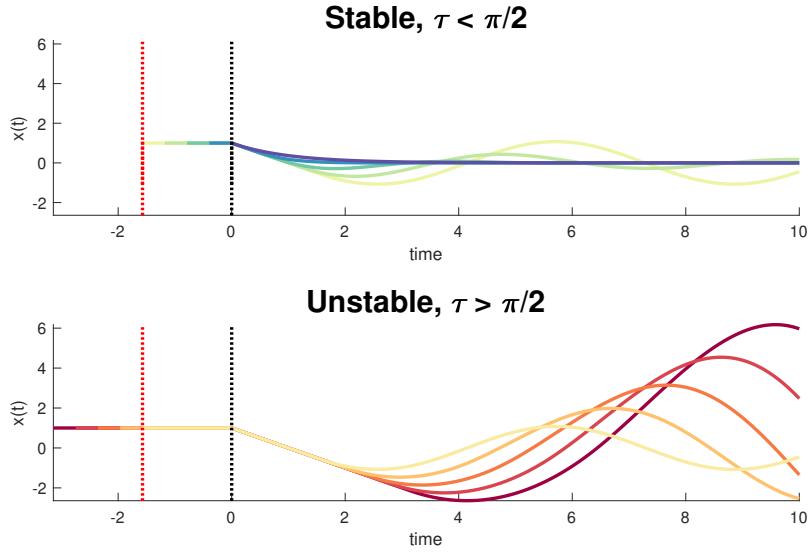


Figure 11.3: Bifurcation of stability as τ exceeds $\pi/2$ in $\dot{x} = -x(t - \tau)$

Problem (11.1) involves a class of histories \mathcal{H} . In this chapter, we will impose that \mathcal{H} is graph-constrained:

Definition 11.2.1. *The history class \mathcal{H} is **graph-constrained** if \mathcal{H} is the set of histories whose graph lies within a given set $H_0 \subseteq [-\tau, 0] \times X$,*

$$\mathcal{H} = \{x_h \in PC([-\tau, 0], X) \mid (t, x_h(t)) \in H_0 \ \forall t \in [-\tau, 0]\},$$

and there are no other continuity restrictions on histories.

11.3 Peak Linear Programs

This section will formulate a primal-dual pair of LPs, each of which upper-bounds Problem (11.1) in objective.

11.3.1 Assumptions

The following assumptions will be imposed on the peak estimation Problem (11.1):

- A1 The sets $\{[-\tau, T], X_0, H_0, X\}$ are all compact with $\tau < T$.
- A2 The function f is Lipschitz inside $[0, T] \times X^2$.
- A3 Any trajectory $x(\cdot \mid x_h)$ with $x_h \in \mathcal{H}$ such that $x(t \mid x_h) \notin X$ for some $t \in [0, T]$ also satisfies $x(t' \mid x_h) \notin X$ for all $t' \geq t$.
- A4 The objective p is continuous.
- A5 The history class \mathcal{H} is graph-constrained by $H_0 \subset [-\tau, 0] \times X$.

In the case where $\tau > T$, the delayed state $t \mapsto x(t - \tau)$ is fully specified in time $[0, T]$ without requiring dynamics information, and (11.1) reduces to a peak estimation problem over ODEs. All tracked histories in \mathcal{H} are bounded due to assumption A1 (since the range X is compact). The nonreturn assumption A3 ensures that a trajectory cannot leave and then return to X to produce a lower value of p , given that the occupation-measure-based techniques used in this chapter can only track trajectories while they are in X .

11.3.2 Measure-Valued Solution

The initial set X_0 is the $t = 0^+$ slice of H_0 . Equation (11.3) describes the measures $(\mu_h, \mu_0, \mu_p, \bar{\mu}_0, \bar{\mu}_1, \nu)$ that will be used to form a free-terminal-time MV-solution to the DDE (11.1b)

with multiple histories (in \mathcal{H}):

$$\text{History} \quad \mu_h \in \mathcal{M}_+(H_0) \quad (11.3a)$$

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (11.3b)$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (11.3c)$$

$$\text{Time-Slack} \quad \nu \in \mathcal{M}_+([0, T] \times X) \quad (11.3d)$$

$$\text{Occupation Start} \quad \bar{\mu}_0 \in \mathcal{M}_+([0, T - \tau] \times X^2) \quad (11.3e)$$

$$\text{Occupation End} \quad \bar{\mu}_1 \in \mathcal{M}_+([T - \tau, T] \times X^2). \quad (11.3f)$$

The joint (relaxed) occupation measure $\bar{\mu} \in \mathcal{M}_+([0, T] \times X^2)$ is constructed from the sum $\bar{\mu} = \bar{\mu}_0 + \bar{\mu}_1$. An MV solution to the DDE in (11.1b) is a set of measures from (11.3) that satisfy three types of constraints: History-Validity, Liouville, Consistency.

11.3.2.1 History-Validity

The first History-Validity constraint is that μ_0 should be a probability distribution over the initial state condition (at $t = 0$). The second is that the history measure μ_h should represent an averaged occupation measure of histories that are defined between $[-\tau, 0]$, which implies that the t -marginal of μ_h should be Lebesgue-distributed. The two History-Validity constraints are

$$\langle 1, \mu_0 \rangle = 1, \quad \pi_{\#}^t \mu_h = \lambda_{[-\tau, 0]}. \quad (11.4)$$

11.3.2.2 Liouville

The true occupation measure $(t, x_0, x_1) \mapsto \bar{\mu}(t, x_0, x_1)$ has a time t , a current state $x_0 = x(t \mid x_h)$, and an external input $x_1 \in X$ with $x_1(t) = x(t - \tau \mid x_h)$. Use of the Liouville equation in (2.8) applied to the joint occupation measure $\bar{\mu} = \bar{\mu}_0 + \bar{\mu}_1$ leads to

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_{\#}^{tx_0} \mathcal{L}_f^\dagger (\bar{\mu}_0 + \bar{\mu}_1). \quad (11.5)$$

11.3.2.3 Consistency

The x_1 input of f from the Liouville equation (11.5) is not arbitrary; it should be equal to a time-delayed $x_1(t) = x_0(t - \tau)$. This requirement will be imposed by a Consistency constraint.

Lemma 11.3.1. *Let $x(\cdot)$ be a solution to (11.1b) for some history x_h with an initial time of 0 and a stopping time of $t^* \in [0, T]$. Then the following two integrals are equal for all $\phi \in C([0, T] \times X)$:*

$$\begin{aligned} & \left(\int_0^{t^*} + \int_{t^*}^{\min(T, t^* + \tau)} \right) \phi(t, x(t - \tau)) dt \\ &= \left(\int_{-\tau}^0 + \int_0^{\min(t^*, T - \tau)} \right) \phi(t' + \tau, x(t)) dt'. \end{aligned} \quad (11.6)$$

Proof. This follows from a change of variable with $t' \leftarrow t - \tau$. \square

Equation (11.6) inspires a consistency constraint for the free-terminal-time MV-solution in (11.3). The left-hand-side of (11.6) may be generalized to

$$\langle \phi(t, x_1), \bar{\mu}_0(t, x_0, x_1) + \bar{\mu}_1(t, x_0, x_1) \rangle + \langle \phi(t, x), \nu(t, x) \rangle, \quad (11.7)$$

in which $\bar{\mu}_0$ is supported in times $[0, \min(t^*, T - \tau)]$, $\bar{\mu}_1$ is supported in times $[T - \tau, t^*]$ if $t^* > T - \tau$, and the slack measure ν implements the $[t^*, \min(T, t^* + \tau)]$ limits. The right-hand-side of (11.6) may be interpreted as

$$\langle \phi(t + \tau, x), \mu_h(t, x) \rangle + \langle \phi(t + \tau, x_0), \bar{\mu}_0(t, x_0, x_1) \rangle. \quad (11.8)$$

Define S^τ as the shift operator $S^\tau \phi(t, x) = \phi(t + \tau, x)$. With an abuse of notation, the pushforward operation $S_\#^\tau$ applied to a measure (such as μ_h) will have the expression

$$\langle \phi, S_\#^\tau \mu_h \rangle = \langle S^\tau \phi, \mu_h \rangle = \langle \phi(t + \tau, x), \mu_h(t, x) \rangle. \quad (11.9)$$

The Consistency constraint inspired by Lemma 11.3.1 is

$$\pi_\#^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) + \nu = S_\#^\tau(\mu_h + \pi_\#^{tx_0}\bar{\mu}_0). \quad (11.10)$$

Remark 11.3.1. *Equation (11.10) may also be written as $\pi_\#^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) \leq S_\#^\tau(\mu_h + \pi_\#^{tx_0}\bar{\mu}_0)$. The associated slack measure is ν .*

11.3.3 Measure Program

An infinite-dimensional LP in terms of the measures from (11.3) to upper-bound Problem (11.1) is

$$p^* = \sup \quad \langle p, \mu_p \rangle \quad (11.11a)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (11.11b)$$

$$\pi_\#^t \mu_h = \lambda_{[-\tau, 0]} \quad (11.11c)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx_0} \mathcal{L}_f^\dagger (\bar{\mu}_0 + \bar{\mu}_1) \quad (11.11d)$$

$$\pi_\#^{tx_1} (\bar{\mu}_0 + \bar{\mu}_1) + \nu = S_\#^\tau (\mu_h + \pi_\#^{tx_0} \bar{\mu}_0) \quad (11.11e)$$

$$\text{Measure Definitions from (11.3).} \quad (11.11f)$$

Remark 11.3.2. *Membership in the history class \mathcal{H} is imposed by the History-Validity constraint (11.11c) and through support of μ_h in (11.3a).*

Definition 11.3.1. *An MV-solution to the DDE (11.1b) with free-terminal-time and histories in \mathcal{H} is a tuple of measures that satisfy (11.11b)-(11.11f) and (11.3a)-(11.3d).*

Theorem 11.3.2. *Under assumptions A1-A5, (11.11) will upper bound (11.1) with $p^* \geq P^*$ when \mathcal{H} is graph-constrained.*

Proof. This proof will proceed by demonstrating that every (t^*, x_h) candidate from (11.1) may be expressed by a unique MV-solution from Defn. 11.3.1. The history measure μ_h is the $[-\tau, 0]$ occupation measure of $x(t)$, and the initial measure μ_0 is the Dirac-delta $\delta_{x_h(0+)}$. The peak measure μ_p is the Dirac-delta $\delta_{t=t^*} \otimes \delta_{x=x(t^*|x_h)}$. The relaxed occupation measures $(\bar{\mu}_0, \bar{\mu}_1, \nu)$ will now be considered. For convenience, define $z(t) = (t, x(t | x_h), x(t - \tau | x_h))$ as the delay embedding of the trajectory $x(t | x_h)$. In the case where $t^* \in [0, T - \tau]$, then $\bar{\mu}_0$ is the $[0, t^*]$ occupation measure of $z(t)$, $\bar{\mu}_1$ is the zero measure, and ν is the $[t^*, t^* + \tau]$ occupation measure of $(t, x(t - \tau | x_h))$. Alternatively when $t^* \in (T - \tau, T]$, $\bar{\mu}_0$ is the $[0, T - \tau]$ occupation measure of $z(t)$, $\bar{\mu}_1$ is the $[T - \tau, t^*]$ occupation measure of $z(t)$, and ν is the $[t^*, T]$ occupation measure of $(t, x(t - \tau | x_h))$. All of the measures in (11.3) have been defined for each input (t^*, x_h) , which proves that $p^* \geq P^*$. \square

Appendix A.9 uses these methods to form MV-solutions to systems with other delay structures (proportional delay, long-delay discrete-time systems).

CHAPTER 11. PEAK ESTIMATION FOR TIME-DELAY SYSTEMS

Remark 11.3.3. *The proof of Theorem 11.3.2 provides a unique MV solution for each DDE trajectory. Additionally, each DDE trajectory given an initial condition x_h is unique under the Lipschitz assumption A2.*

We note that MV solutions are not necessarily unique (for a given terminal time distribution) when the history measure μ_h is supported on the graph of more than one curve. As an example, Figure 11.4 shows two sets of curves under the dynamics $\dot{x}(t) = -2x(t) - 3x(t-1)$ in the times $t \in [0, 5]$. The history occupation measure $\mu_h = 0.5\lambda_{[-1,0]} \otimes \delta_{x=1} + 0.5\lambda_{[-1,0]} \otimes \delta_{x=-1}$ is supported in the set $\mu_h \in [-1, 0] \times \{-1, 1\}$. The superposition of each set of red and blue curves each have the same history measure μ_h , but the switch that takes place on the bottom plot (e.g. blue: $x_h(t) = 1$ for $t \in [-1, -0.5]$, $x_h(t) = -1$ for $t \in [0.5, 0]$) yields a different trajectory going forward in time.

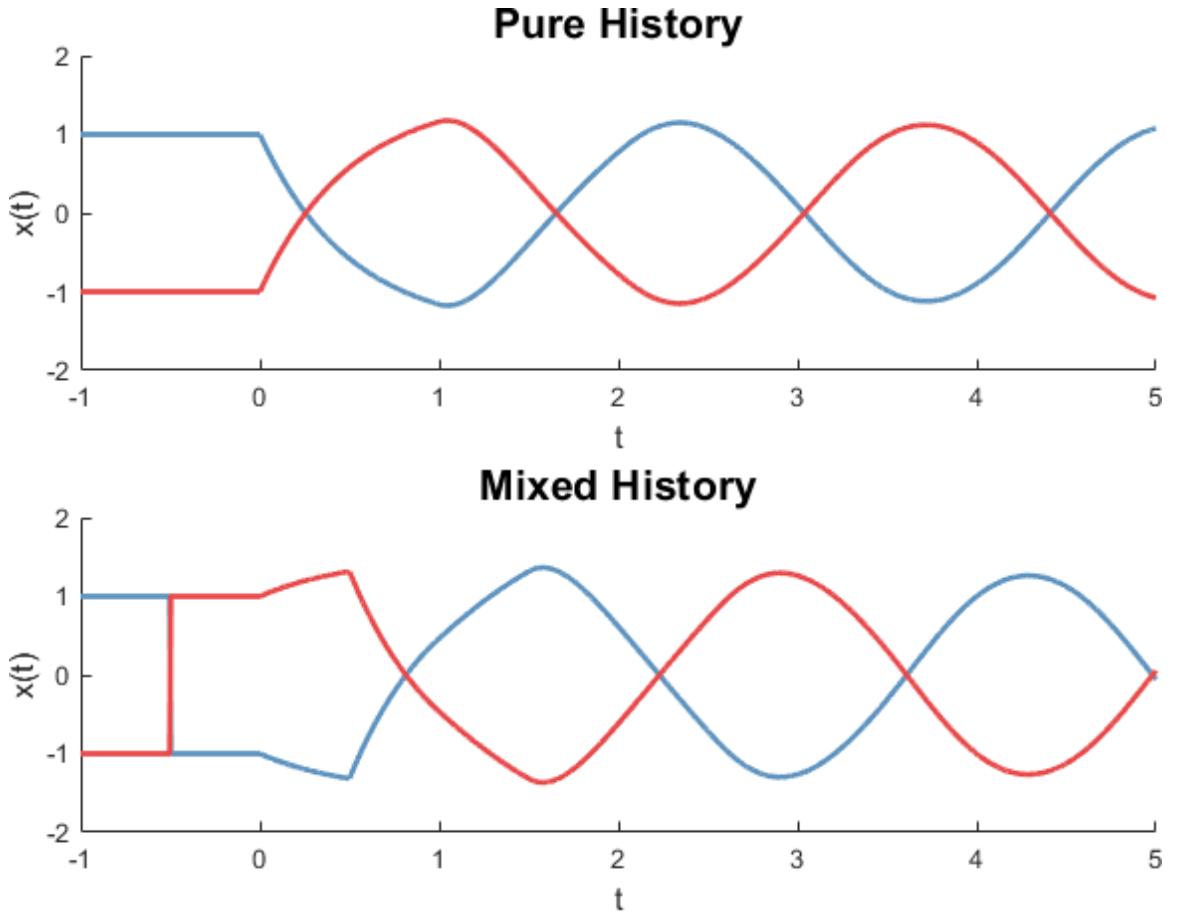


Figure 11.4: The same μ_h leads to different trajectories in times $(0, 5]$

11.3.4 Function Program

The dual program of (11.11) with variables (γ, ξ, v, ϕ) is

$$d^* = \inf_{\gamma \in \mathbb{R}} \gamma + \int_{-\tau}^0 \xi(t) dt \quad (11.12a)$$

$$\xi(t) + \phi(t + \tau, x) \geq 0 \quad \forall (t, x) \in H_0 \quad (11.12b)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (11.12c)$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (11.12d)$$

$$\phi(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (11.12e)$$

$$\mathcal{L}_f v(t, x_0) + \phi(t, x_1) \leq \phi(t + \tau, x_0) \quad \forall (t, x_0, x_1) \in [0, T - \tau] \times X^2 \quad (11.12f)$$

$$\mathcal{L}_f v(t, x_0) + \phi(t, x_1) \leq 0 \quad \forall (t, x_0, x_1) \in [T - \tau, T] \times X^2 \quad (11.12g)$$

$$\xi \in C([-\tau, 0]) \quad (11.12h)$$

$$v \in C^1([0, T] \times X) \quad (11.12i)$$

$$\phi \in C([0, T] \times X). \quad (11.12j)$$

Theorem 11.3.3. *There is no duality gap between (11.11) and (11.12).*

Proof. See Appendix A.10. □

We pose the following conjecture based on [7, 161]:

Conjecture 11.3.1. *Assume that A1-A5 hold. Additionally, assume that $T > \tau > 0$ and the image-set $f(t, x_0, X)$ is convex for all fixed $t \in [0, T]$, $x_0 \in X$. Then there is no relaxation gap between (11.1) and (11.11) ($p^* = P^*$).*

Proving Conjecture 11.3.1 is the subject of ongoing work.

Appendix A.11 contains a further discussion of the continuity and structural aspects of the dual solution in (11.12) as applied to bounding costs on DDE OCPs.

11.4 Peak Moment Program

This section will briefly review the moment-SOS hierarchy [27] in order to approximate-from-above Program (11.11) by a sequence of finite-dimensional SDPs.

11.4.1 Moment Program

Additional assumptions are required in order to approximate (11.11) using the moment-SOS hierarchy:

A6 The sets H_0 , X_0 , and X are Archimedean BSA sets;

A7 Both p and f are polynomials.

Let the measures $(\mu_h, \mu_0, \mu_p, \bar{\mu}_0, \bar{\mu}_1, \nu)$ have associated pseudo-moment sequences $(\mathbf{m}^h, \mathbf{m}^0, \mathbf{m}^p, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1, \mathbf{m}^\nu)$ respectively. Let $\alpha \in \mathbb{N}^n$ and $\beta \in \mathbb{N}$ be multi-indices that define monomial test functions $x_0^\alpha t^\beta$. For each multi-index tuple (α, β) , the operator $\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}^p, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1)$ may be derived from the linear relations induced by the Liouville equation (11.11d) (in which $\delta_{\beta 0} = 1$ is a Kronecker delta):

$$0 = \langle x^\alpha, \mu_0 \rangle \delta_{\beta 0} + \langle \mathcal{L}(x_0^\alpha t^\beta), \bar{\mu}^0 + \bar{\mu}^1 \rangle - \langle x^\alpha t^\beta, \mu_\tau \rangle. \quad (11.13)$$

Similarly, the operator $\text{Cons}_{\alpha\beta}(\mathbf{m}^h, \mathbf{m}^\nu, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1)$ may be derived from the consistency constraint (11.11e) by

$$\begin{aligned} 0 = & \langle x_1^\alpha t^\beta, \bar{\mu}^0 + \bar{\mu}^1 \rangle + \langle x^\alpha t^\beta, \nu \rangle - \langle x^\alpha (t + \tau)^\beta, \mu_h \rangle \\ & - \langle x_0^\alpha (t + \tau)^\beta, \bar{\mu}^0 \rangle. \end{aligned} \quad (11.14)$$

Given a degree $d \in \mathbb{N}$, the dynamics degree $\tilde{d} \geq d$ may be defined as $\tilde{d} = d + \lfloor \deg f / 2 \rfloor$.

Problem 11.4.1. *Program (11.11) is upper-bounded by the following order-d LMI in pseudo-moments:*

$$p_d^* = \max \sum_{\alpha \in \mathbb{N}^n} p_\alpha \mathbf{m}_\alpha^p \quad (11.15a)$$

$$\mathbf{m}_0^0 = 1 \quad (11.15b)$$

$$\forall (\alpha, \beta) \in \mathbb{N}_{\leq 2d}^{n+1} :$$

$$\mathbf{m}_\beta^h = \int_{-\tau}^0 t^\beta dt = -(-\tau)^{\beta+1}/(\beta+1) \quad (11.15c)$$

$$\text{Liou}_{\alpha\beta}(\mathbf{m}^0, \mathbf{m}^p, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1) = 0 \quad (11.15d)$$

$$\text{Cons}_{\alpha\beta}(\mathbf{m}^h, \mathbf{m}^\nu, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1) = 0 \quad (11.15e)$$

$$\mathbb{M}_d((X_0)\mathbf{m}^0), \mathbb{M}_{\tilde{d}}((H_0)\mathbf{m}^h) \succeq 0 \quad (11.15f)$$

$$\mathbb{M}_d(([0, T] \times X)\mathbf{m}^p), \mathbb{M}_{\tilde{d}}(([0, T] \times X)\mathbf{m}^\nu) \succeq 0 \succeq 0 \quad (11.15g)$$

$$\mathbb{M}_{\tilde{d}}(([0, T - \tau] \times X^2)\bar{\mathbf{m}}^0) \succeq 0 \quad (11.15h)$$

$$\mathbb{M}_{\tilde{d}}(([T - \tau, T] \times X^2)\bar{\mathbf{m}}^1) \succeq 0. \quad (11.15i)$$

CHAPTER 11. PEAK ESTIMATION FOR TIME-DELAY SYSTEMS

The objective (11.15a) is the pseudo-moment version of $\langle p, \mu_p \rangle$. Constraints (11.15c) and (11.15b) are History-Validity constraints from (11.4) when applied to the pseudo-moments $(\mathbf{m}^\nu, \mathbf{m}^0)$. Constraints (11.15d) and (11.15e) are the Liouville and Consistency constraints respectively. Constraints (11.15f)-(11.15i) are support constraints necessary for the pseudo-moments to have representing measures.

Boundedness of all moments of measures in (11.3) is required to obtain convergence of (11.15) to (11.11) as $d \rightarrow \infty$.

Lemma 11.4.2. *All measures from (11.3) in an MV-solution (Defn. 11.3.1) are bounded under assumptions A1-A7.*

Proof. Boundedness of a measure's mass and support is a sufficient condition that all of the measure's moments are bounded. Assumption A1 ensures compactness, with the requirement from Defn. 11.2.1 that $H_0 \subseteq [-\tau, X]$ and $X_0 \subseteq X$. The remainder of this proof will involve finding upper bounds on the masses of all measures in (11.3).

The initial measure μ_0 has a mass of 1 and the history measure μ_h has a mass of τ by the History-Validity constraints (11.11b) and (11.11c). Substitution of the test function $v(t, x) = 1$ in the Liouville (11.11d) leads to $\langle 1, \mu_p \rangle = \langle 1, \mu_0 \rangle = 1$. Since T is finite, the moment $\langle t, \mu_p \rangle \leq \langle 1, \mu_p \rangle (\sup_{t \in [0, T]} t) = T$ is also finite. Use of the test function $v(t, x) = t$ into the Liouville (11.11d) yields $\langle t, \mu_p \rangle = \langle 1, \bar{\mu}_0 + \bar{\mu}_1 \rangle \leq T$. Because $\bar{\mu}_0$ and $\bar{\mu}_1$ are both nonnegative Borel measures, it holds that $\langle 1, \bar{\mu}_0 \rangle \leq T$ and $\langle 1, \bar{\mu}_1 \rangle \leq T$. The final constraint involves substitution of $\phi(t, x) = 1$ into the Consistency (11.11e), resulting in

$$\begin{aligned} \langle 1, \bar{\mu}_0 + \bar{\mu}_1 \rangle + \langle 1, \nu \rangle &= \langle 1, \mu_h \rangle + \langle 1, \bar{\mu}_0 \rangle \\ \langle 1, \nu \rangle &= \langle 1, \mu_h \rangle - \langle 1, \bar{\mu}_1 \rangle = \tau - \langle 1, \bar{\mu}_1 \rangle. \end{aligned} \tag{11.16}$$

Given that $\bar{\mu}_1$ and ν are nonnegative Borel measures and cannot have negative masses, the mass $\langle 1, \nu \rangle$ is constrained within $[0, \tau]$. All masses are demonstrated to be finite, thus proving boundedness. \square

Remark 11.4.1. *Neglecting the History-Validity constraint (11.11c) allows for μ_h in (11.16) to have infinite mass, violating the boundedness principle.*

Theorem 11.4.3. *The optima in (11.15) will converge as $\lim_{d \rightarrow \infty} p_d^* = p^*$ to (11.11) under assumptions A1-A6*

Proof. This follows from Corollary 8 of [23] under the boundedness condition in Lemma 11.4.2. \square

Remark 11.4.2. Assumption A6 can be generalized to cases where the sets (H_0, X_0, X) are the unions of BSA sets. As an example, consider $H_0 = H_0^1 \cup H_0^2$ in which $\pi^t H_0^1 = [-\tau, -\tilde{\tau}]$ and $\pi^t H_0^2 = [-\tilde{\tau}, 0]$ for some $\tilde{\tau} \in (0, \tau)$. Then the pseudo-moments $\mathbf{m}^h = \mathbf{m}_1^h + \mathbf{m}_2^h$ can be implicitly constructed from $\mathbb{M}_d((H_0^1)\mathbf{m}_1^h), \mathbb{M}_d((H_0^2)\mathbf{m}_2^h) \succeq 0$.

11.4.2 Computational Complexity

The size of the order- d PSD moment matrices associated with the pseudo-moment sequences $(\mathbf{m}^h, \mathbf{m}^0, \mathbf{m}^p, \bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1, \mathbf{m}^\nu)$ are listed in Table 11.1.

Table 11.1: Size of Moment Matrices in LMI (11.15)

Matrix:	$\mathbb{M}_d(\mathbf{m}^0)$	$\mathbb{M}_{\tilde{d}}(\mathbf{m}^p)$	$\mathbb{M}_d(\mathbf{m}^h)$
Size:	$\binom{n+d}{d}$	$\binom{n+1+d}{d}$	$\binom{n+1+\tilde{d}}{\tilde{d}}$
Matrix:	$\mathbb{M}_d(\bar{\mathbf{m}}^0)$	$\mathbb{M}_{\tilde{d}}(\bar{\mathbf{m}}^1)$	$\mathbb{M}_d(\mathbf{m}^\nu)$
Size:	$\binom{2n+1+\tilde{d}}{\tilde{d}}$	$\binom{2n+1+\tilde{d}}{\tilde{d}}$	$\binom{n+1+\tilde{d}}{\tilde{d}}$

The largest size written in Table 11.1 is $\binom{2n+1+\tilde{d}}{\tilde{d}}$, which occurs with the pseudo-moment sequences $(\bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1)$ associated to the two joint occupation measures $(\bar{\mu}_0, \bar{\mu}_1)$. Equality constraints between entries of the moment matrices must be added to convert the LMI into an SDP for use in symmetric-cone Interior Point Methods. The per-iteration complexity of solving an SDP derived from an order- d LMI involved in the moment-SOS hierarchy scales as $O(n^{6d})$ [27] with n . In the case of LMI (11.15), the computational complexity of solving (11.15) will scale approximately as $(2n + 1)^{6\tilde{d}}$ (based on $\bar{\mathbf{m}}^0, \bar{\mathbf{m}}^1$).

11.5 Extensions

This section discusses several extensions to the DDE peak estimation framework.

11.5.1 Distance Estimation

The distance estimation framework of Chapter 5 may also be applied towards DDEs. The DDE distance estimation program with metric c and unsafe set $X_u \subset X$ is

$$P^* = \inf_{t^* \in [0, T], x_h(\cdot)} c(x(t^* | x_h); X_u) \quad (11.17a)$$

$$\dot{x} = f(t, x(t), x(t - \tau)) \quad \forall t \in [0, T] \quad (11.17b)$$

$$x(t) = x_h(t) \quad \forall t \in [-\tau, 0] \quad (11.17c)$$

$$x_h(\cdot) \in \mathcal{H}. \quad (11.17d)$$

Safety in program (11.17) is measured pointwise: a trajectory is safe if $x(t | x_h) \notin X_u$ for every time $t \in [0, T]$. Safety ensuring that the entire history is never contained within X_u ($\exists s \in [-\tau, 0] | x(t - s | x_h) \notin X_u \forall t \in [0, T]$) is a more challenging separate problem and will not be considered here.

The MV-solution in (11.3) may be applied to (11.17) to create a DDE version of the DDE distance estimation task in (5.11) by adding a joint probability measure η :

$$c^* = \inf \langle c, \eta \rangle \quad (11.18a)$$

$$\pi_\#^x \mu^p = \pi_\#^x \eta \quad (11.18b)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (11.18c)$$

$$\pi_\#^t \mu_h = \lambda_{[-\tau, 0]} \quad (11.18d)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx_0} \mathcal{L}_f^\dagger(\bar{\mu}_0 + \bar{\mu}_1) \quad (11.18e)$$

$$\pi_\#^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) + \nu = S_\#^\tau(\mu_h + \pi_\#^{tx_0} \bar{\mu}_0) \quad (11.18f)$$

$$\eta \in \mathcal{M}_+(X \times X_u) \quad (11.18g)$$

$$\text{Measure Definitions from (11.3).} \quad (11.18h)$$

The distance estimation program only affects the cost (11.18a). This change is orthogonal to the modification in dynamics necessary to create a DDE MV-solution from the ODE program.

11.5.2 Shaping Constraints

Assumption A5 imposes that the class \mathcal{H} is graph-constrained. Some applications involve further structure in the function class \mathcal{H} , such as requiring that the histories in \mathcal{H} are constant in time between $t \in [-\tau_r, 0]$. Examples of these constant histories for the system in (11.2) starting within the black box (H_0) are plotted in Figure 11.5.

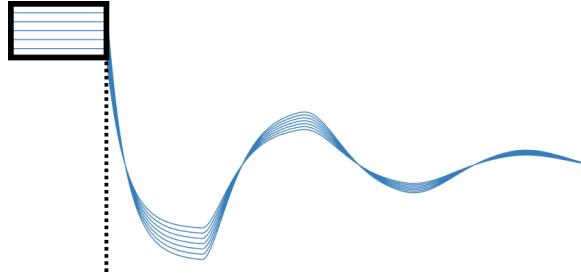


Figure 11.5: Constant histories in the black box

These types of structure in histories may be realized by adding constraints to μ_h . A method to ensure that the histories in μ_h are constant in time between $t \in [-\tau, 0]$ is by requiring μ_h to be the occupation measure of the system $\dot{x} = 0$ through a Liouville equation

$$\langle v(0, x), \mu_0 \rangle = \langle \partial_t v(t, x), \mu_h \rangle + \langle v(-\tau, x), \mu_0 \rangle \quad \forall v \in C([-\tau_r, 0]). \quad (11.19)$$

11.5.3 Multiple Time Delays

A DDE with multiple time-delays $0 < \tau_1 < \tau_2 < \dots < \tau_r$ for (r, τ_r) finite and a history $x_h \in PC([-\tau_r, 0], X)$ is

$$\dot{x}(t) = f(t, x(t), x(t - \tau_1), \dots, x(t - \tau_r)) \quad (11.20)$$

$$x(t) = x_h(t), \quad \forall t \in [-\tau_r, 0].$$

A peak estimation problem for (11.20) with history class \mathcal{H} and objective $p(x)$ is

$$P^* = \sup_{t^* \in [0, T], x_h(\cdot)} p(x(t^* | x_h)) \quad (11.21a)$$

$$\dot{x} = f(t, x(t), x(t - \tau_1), \dots, x(t - \tau_r)) \quad \forall t \in [0, T] \quad (11.21b)$$

$$x(t) = x_h(t) \quad \forall t \in [-\tau_r, 0] \quad (11.21c)$$

$$x_h(\cdot) \in \mathcal{H}. \quad (11.21d)$$

A multiple-time-delay MV-solution for the peak estimation problem (11.21) is (for $i =$

$1..r$):

$$\text{History} \quad \mu_{hi} \in \mathcal{M}_+ (H_0 \cap ([-\tau_i, -\tau_{i-1}] \times X)) \quad (11.22a)$$

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (11.22b)$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (11.22c)$$

$$\text{Time-Slack} \quad \nu_i \in \mathcal{M}_+([0, T] \times X) \quad (11.22d)$$

$$\text{Occupation Start} \quad \bar{\mu}_0 \in \mathcal{M}_+([0, T - \tau] \times X^2) \quad (11.22e)$$

$$\text{Occupation End} \quad \bar{\mu}_i \in \mathcal{M}_+([T - \tau_i, T - \tau_{i-1}] \times X^2). \quad (11.22f)$$

The Lie derivative operator \mathcal{L}_f with respect to (11.20) for $v \in C^1([0, T] \times X)$ is

$$\mathcal{L}_f v(t, x_0) = \partial_t v(t, x_0) + f(t, x_0, x_1, \dots, x_r) \cdot \nabla_{x_0} v(t, x_0). \quad (11.23)$$

The multiple-time-delay peak estimation LP for (11.21) problem of $p(x)$ is

$$p^* = \sup \langle p, \mu_p \rangle \quad (11.24a)$$

$$\langle 1, \mu_0 \rangle = 1 \quad (11.24b)$$

$$\pi_\#^t \mu_{hi} = \lambda_{[-\tau_i, -\tau_{i-1}]} \quad \forall i = 1..r \quad (11.24c)$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx_0} \mathcal{L}_f^\dagger (\bar{\mu}_0 + \sum_{i=1}^r \bar{\mu}_i) \quad (11.24d)$$

$$\pi_\#^{tx_1} (\bar{\mu}_0 + \sum_{i=1}^r \bar{\mu}_i) + \nu_i = S_\#^{\tau_i} (\sum_{j=1}^i \mu_{hj} + \pi_\#^{tx_0} (\bar{\mu}_0 + \sum_{j=1}^{i-1} \bar{\mu}_j)) \quad \forall i = 1..r \quad (11.24e)$$

$$\text{Measure Definitions from (11.22).} \quad (11.24f)$$

Theorem 11.3.2 can be extended to the multiple-time-delay case to prove that $P^* \leq p^*$ between (11.21) and (11.24). Even if Conjecture 11.3.1 holds in the single-delay case, it is unlikely the conjecture is satisfied in the multiple-delay case due to findings in [163].

11.5.4 Uncertainty

This extension subsection will discuss three types of uncertainty that can affect DDEs dynamics: time-independent, time-dependent, and unknown-delay.

11.5.4.1 Time-independent Uncertainty

Time-independent uncertainty $\theta \in \Theta$ for a set Θ can be added to dynamics by adjoining the state θ following $\dot{\theta} = 0$ to (11.1b). This same process occurs in Section 4.2.1 for the ODE case.

11.5.4.2 Time-dependent Uncertainty

Time-dependent uncertainty may be implemented by a Young Measure approach as in Section 4.2.2. Given dynamics $\dot{x}(t) = f(t, x(t), x(t - \tau), w(t))$ for $w(t) \in W$, the joint occupation measures representing trajectories are $\bar{\mu}_0 \in \mathcal{M}_+([0, T - \tau] \times X^2 \times W)$ and $\bar{\mu}_1 \in \mathcal{M}_+([T - \tau, T] \times X^2 \times W)$. No substantial changes are required to the Liouville nor consistency constraints.

Input delays may also be introduced into dynamics with $\dot{x}(t) = f(t, x(t), x(t - \tau), w(t), w(t - \tau))$ under the state history x_h and input history w_h (defined in times $t \in [-\tau, 0]$). The associated joint occupation measures are now $\bar{\mu}_0 \in \mathcal{M}_+([0, T - \tau] \times X^2 \times W^2)$ and $\bar{\mu}_1 \in \mathcal{M}_+([T - \tau, T] \times X^2 \times W^2)$, each involving variables (t, x_0, x_1, w_0, w_1) . The state-input history class is $\mathcal{H} \in PC([-\tau, 0], X \times W)$, and its history occupation measure now includes an input component $\mu_h \in \mathcal{M}_+([-\tau, 0] \times X \times W)$.

While the Liouville equation stays the same as (11.5), the consistency constraint ensures that the w_1 coordinate contains a delayed copy of w_0 with

$$\pi_{\#}^{tx_1w_1}(\bar{\mu}_0 + \bar{\mu}_1) = S_{\#}^{\tau}(\mu_h + \pi^{tx_0w_0}\bar{\mu}_0). \quad (11.25)$$

11.5.4.3 Unknown Delays

This extension focuses on dynamics where the time-independent (constant) delay is unknown but fixed in the finite range of $\tau \in [\underline{\tau}, \bar{\tau}]$. The unknown delay τ must be treated as an additional state with $\dot{\tau} = 0$.

The following support sets may be defined:

$$\Omega_h = \{(\tau, t, x) \mid \tau \in [\underline{\tau}, \bar{\tau}], (t, x) \in H_0 \mid_{\tau}\} \quad (11.26a)$$

$$\Omega_0 = \{(\tau, t, x_0, x_1) \mid \tau \in [\underline{\tau}, \bar{\tau}], t \in [0, T - \tau], (x_0, x_1) \in X^2\} \quad (11.26b)$$

$$\Omega_1 = \{(\tau, t, x_0, x_1) \mid \tau \in [\underline{\tau}, \bar{\tau}], t \in [T - \tau, T], (x_0, x_1) \in X^2\}. \quad (11.26c)$$

An MV-solution in the unknown-delay case has the form

$$\text{History} \quad \mu_h \in \mathcal{M}_+(\Omega_h) \quad (11.27\text{a})$$

$$\text{History Slack} \quad \bar{\mu}_h \in \mathcal{M}_+(\Omega_h) \quad (11.27\text{b})$$

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0 \times [\underline{\tau}, \bar{\tau}]) \quad (11.27\text{c})$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X \times [\underline{\tau}, \bar{\tau}]) \quad (11.27\text{d})$$

$$\text{Time-Slack} \quad \nu \in \mathcal{M}_+([0, T] \times X \times [\underline{\tau}, \bar{\tau}]) \quad (11.27\text{e})$$

$$\text{Occupation Start} \quad \bar{\mu}_0 \in \mathcal{M}_+(\Omega_0) \quad (11.27\text{f})$$

$$\text{Occupation End} \quad \bar{\mu}_1 \in \mathcal{M}_+(\Omega_1). \quad (11.27\text{g})$$

The Liouville and Consistency constraints in the unknown-delay case are unchanged as compared to the known-delay system (with the new state $\dot{\tau} = 0$).

However, the history-validity constraints have the following form:

$$\langle 1, \mu_0 \rangle = 1 \quad (11.28\text{a})$$

$$\delta_{t=0} \otimes (\pi_\#^\tau \mu_0) = \delta_{t=-\bar{\tau}} \otimes (\pi_\#^\tau \mu_0) + (\partial_t)_\# (\pi_\#^{t\tau} (\mu_h + \hat{\mu}_h)) \quad (11.28\text{b})$$

$$\pi^t (\mu_h + \hat{\mu}_h) = \lambda_{[-\bar{\tau}, 0]}. \quad (11.28\text{c})$$

Constraint (11.28a) ensures that the initial distribution μ_0 is a probability measure. Constraint (11.28b) imposes that $\tau(t)$ is constant in time between $t = [-\bar{\tau}, 0]$. Constraint (11.28c) is a domination term that requires the history x_h to be defined in times $[-\bar{\tau}, 0]$.

It is an open problem to extend consistency constraints and MV-solutions towards cases where the delay $\tau(t)$ is time-dependent (such as $\dot{\tau}(t) \in [-B, B]$).

11.6 Numerical Examples

All experiments were developed in MATLAB 2021a, and code is available at <https://github.com/Jarmill/timedelay>. Dependencies include Gloptipoly [30], YALMIP [48], and Mosek [49] in order to formulate and solve moment-SOS LMIs and SDPs.

In this section, a notational convention where (x_1, x_2) correspond to coordinates of $x \in X$ will be used. All sampled histories in visualizations are piecewise-constant inside H_0 with 10 randomly-spaced jumps between $[-\tau, 0]$.

11.6.1 Epidemic Model

This section provides an example of a MV-solution and peak estimation given a single history in a compartmental epidemic model. Many diseases have incubation periods during which there is a delay between initial infection and infectious potential. In the current COVID-19 pandemic, this incubation period appears to be between 2-14 days, with a median of 5 days [173]. The epidemic dynamics with time delays are

$$S'(t) = -\beta S(t)I(t) \quad (11.29a)$$

$$I'(t) = \beta S(t-\tau)I(t-\tau) - \gamma I(t) \quad (11.29b)$$

$$R'(t) = \gamma I(t). \quad (11.29c)$$

There exists also exists a ‘latent’ state $L'(t) = \beta S(t)I(t) - \beta S(t-\tau)I(t-\tau)$ such that $S+I+R+L = 1$. The setting discussed in this section is $\beta = 0.4$, $\gamma = 0.1$, $T = 30$.

Figures 11.6a and 11.6b display simulations of this epidemic model as τ changes under a constant state history with $R = 0$. The black curve in Figures 11.6a and 11.6b is the plot of $I(t)$ at $\tau = 0$. As the incubation period τ , the time t^* at which the peak is achieved is delayed (moves rightwards) in a monotonically increasing manner. The other colored curves in each plot have delays $\tau \in 1..9$. In Figure 11.6a with $I_h = 0.1$, the peak infected population decreases as the delay τ increases. Conversely in Figure 11.6b with $I_h = 0.2$, the peak infected population increases as the delay increases.

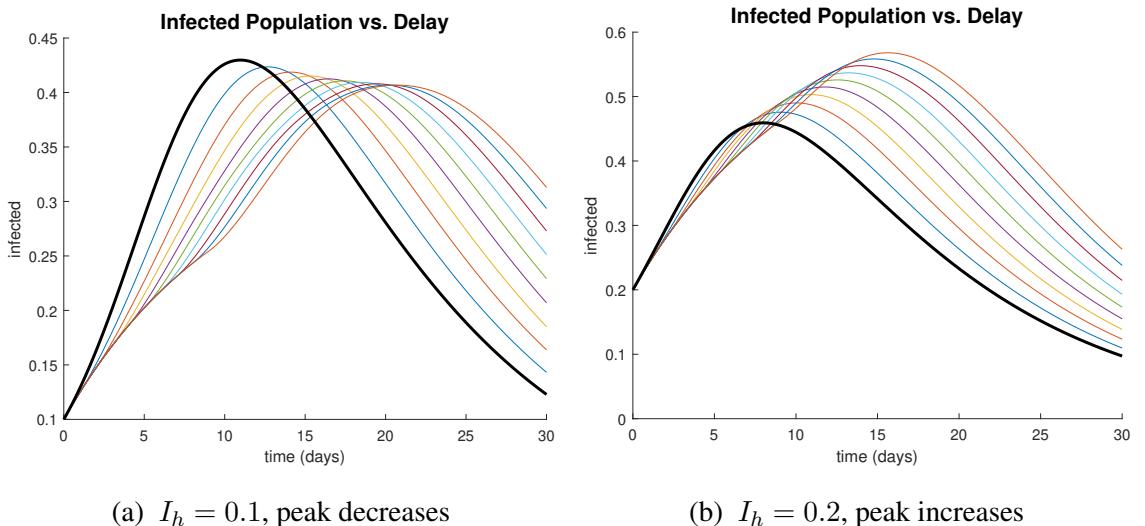


Figure 11.6: Peak infected population vs. time delay

For the peak estimation example, a constant history is assumed with an initial infection rate of $I_h = 0.2$ and an incubation period of $\tau = 9$, forming the initial history $S(t) = 1 - I_h, I(t) = I_h, R(t) = 0$ between $t \in [-9, 0]$.

For numerical purposes the dynamics are scaled such that $\tilde{t} \in [0, 1]$ with an effective delay of $\tilde{\tau} = \tau/T = 0.3$. Only the $x = (S, I)$ subsystem is considered to form the state set $X = \{S \geq 0, I \geq 0, S + I \leq 1\}$, and the joint occupation measures $(\bar{\mu}_0, \bar{\mu}_1)$ have variables (t, S_0, I_0, S_1, I_1) .

Peak estimation is employed to bound the maximum infection rate over the course of the epidemic. This peak estimation program maximizes $\langle I, \mu_p \rangle$ under the constraint that $(\mu_p, \bar{\mu}, \{\nu_0, \nu_1\}, \hat{\nu}_1)$ is a free-time MV-solution of dynamics (11.29).

Figure 11.7 displays the output of peak estimation, where the order-3 LMI relaxation bounds the maximal infection rate at 56.89%. The moment matrix $\mathbb{M}_3[y_p]$ is approximately rank-1 (second largest eigenvalue of $\mathbb{M}_3[y_p] = 2.448 \times 10^{-5}$), and the extracted optimum from $\mathbb{M}_3[y_p]$ by Algorithm 1 is $(S^*, I^*) = (0.0561, 0.5689)$ occurring at $t^* = 15.636$ days.

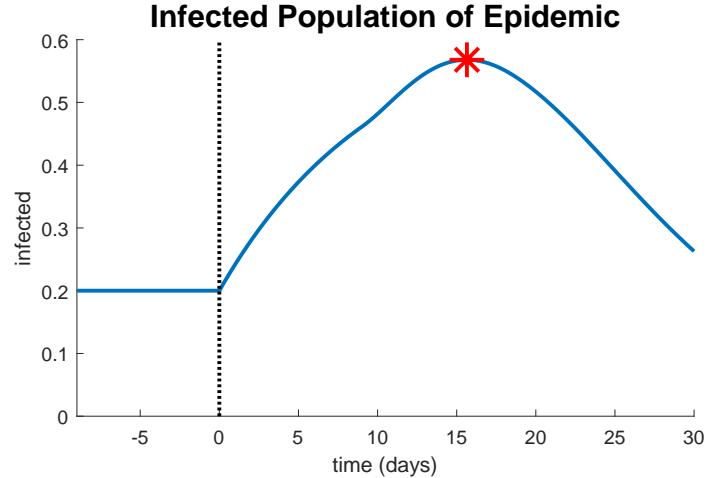


Figure 11.7: SIR peak estimation and recovery at order 3

11.6.2 Delayed Flow System

A time-delayed version of the Flow system from [36] is

$$\dot{x}(t) = \begin{bmatrix} x_2(t) \\ -x_1(t - \tau) - x_2(t) + x_1(t)^3/3 \end{bmatrix}. \quad (11.30)$$

Figure 11.8 plots the delayed Flow system (11.30) without lag ($\tau = 0$ in blue) and with a lag ($\tau = 0.75$ in orange) starting from the constant initial history $x_h(t) = (1.5, 0)$, $\forall t \in [-\tau, 0]$ (black circle).

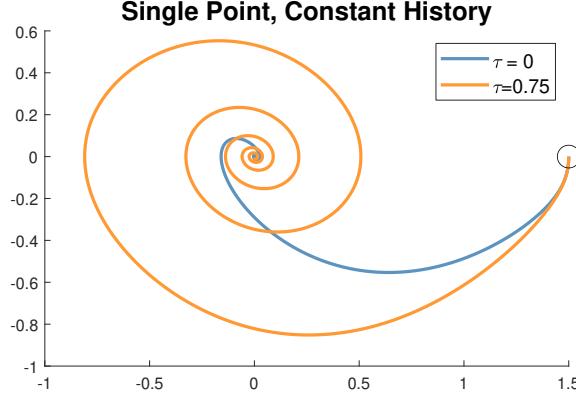
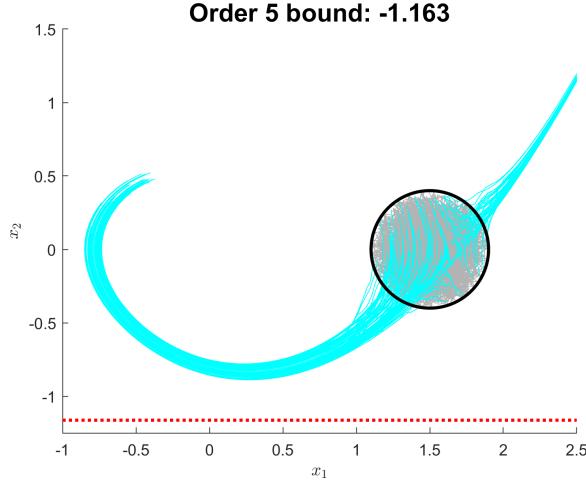


Figure 11.8: Comparison of delayed Flow systems (11.30) with lags $\tau = 0$ and $\tau = 0.75$ in times $t \in [0, 20]$

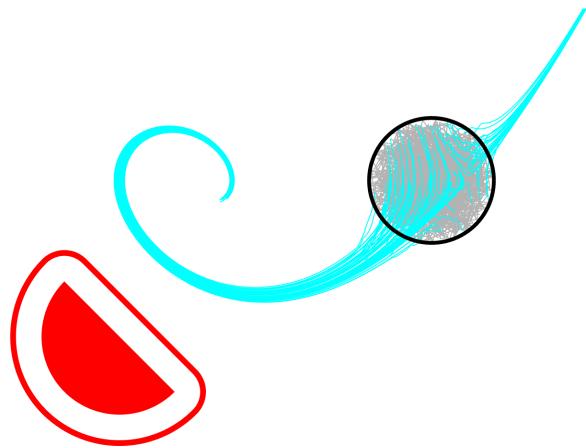
The time-zero set of allowable histories is $X_0 = \{x \in \mathbb{R}^2 \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$. The history class \mathcal{H} will be the set of functions $x_h \in PC([-0.75, 0])$ whose graphs $(t, x(t))$ are contained within the cylinder $H_0 = [-0.75, 0] \times X_0$. No further requirements of continuity are posed on histories in \mathcal{H} . The considered peak estimation aims to find the minimum value of x_2 (maximize $p(x) = -x_2$) for trajectories following (11.30) starting from H_0 , within the state set $X = [-1.25, 2.5] \times [-1.25, 1.5]$ and time horizon $T = 5$. The first five bounds on the maximum value of $-x_2$ by solving (11.15) are $p_{1:5}^* = [1.25, 1.2183, 1.1913, 1.1727, 1.1630]$.

Figure 11.9 plots trajectories and peak information associated with this example. The black circle is the initial set X_0 . The initial histories inside X_0 are plotted in grey. These sampled histories are piecewise constant with 10 uniformly spaced jumps (moving to a new point uniformly sampled in X_0) within $[-0.75, 0]$. The cyan curves are the DDE trajectories of (11.30) starting from the grey histories. The red dotted line is the p_5^* bound on the minimum vertical coordinate of a point on any trajectory starting from \mathcal{H} up to $T = 5$.

Distance estimation is performed on the Flow system 11.30 with an L_2 metric, a time horizon of $T = 8$, arbitrarily varying histories in H_0 , a time horizon of $\tau = 0.5$, and a half-circle unsafe set $X_u = \{x \mid 0.5^2 \geq (x_1 + 0.5)^2 + (x_2 + 1)^2, (1.5 + x_1 + x_2)\}$. The recovered distance estimates up to degree 4 from SDP relaxations of (11.18) are $c_{1:4}^* = [1.1897 \times 10^{-4}, 4.0420 \times 10^{-4}, 0.1572, 0.1820]$. Figure 11.10 plots the set X_u in red along with its $c_4^* = 0.1820$ certified


 Figure 11.9: Minimize x_2 on the delayed Flow system (11.30)

distance contour.


 Figure 11.10: Minimize $c(x; X_u)$ on the delayed Flow system (11.30)

11.6.3 Delayed Time-Varying System

This example involves peak estimation of a DDE version of the time-varying Example 2.1 of [6] with

$$\dot{x}(t) = \begin{bmatrix} x_2(t)t - 0.1x_1(t) - x_1(t-\tau)x_2(t-\tau) \\ -x_1(t)t - x_2(t) + x_1(t)x_1(t-\tau) \end{bmatrix}. \quad (11.31)$$

The considered support parameters are $\tau = 0.75$, $T = 5$, and $X = [-1.25, 1.25] \times [-0.75, 1.25]$. The time-zero set is the disk $X_0 = \{x \in \mathbb{R}^2 \mid (x_1 + 0.75)^2 + x_2^2 \leq 0.3^2\}$.

CHAPTER 11. PEAK ESTIMATION FOR TIME-DELAY SYSTEMS

The only restriction on allowable histories \mathcal{H} is that their graphs are contained in the history set $H_0 = [-0.75, 0] \times X_0$.

Solving the SDP associated with the LMI (11.15) to maximize the peak function $p = x_1$ yields the sequence of five bounds $p_{1:5}^* = [1.25, 1.25, 1.1978, 0.8543, 0.718264618]$. Figure 11.11 plots system trajectories and the p_5^* bound on x_1 using the same visual convention as Figure 11.9 (black circle X_0 , grey histories $x_h(t)$, cyan trajectories $x(t | x_h)$, red dotted line $x_1 = p_5^*$).

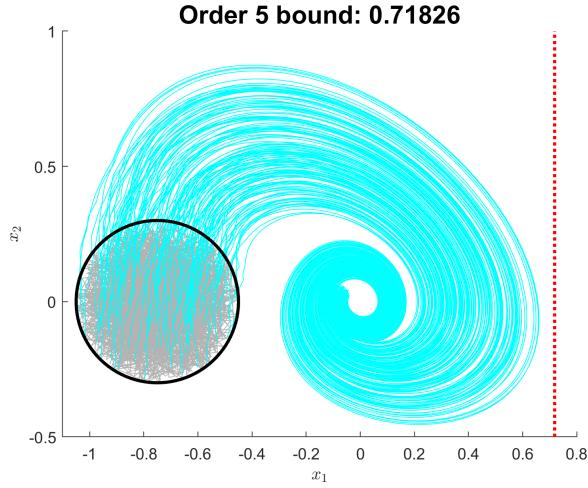


Figure 11.11: Maximize x_1 on the delayed time-varying (11.31)

Figure 11.12 plots the corresponding trajectory and bound information in 3d (t, x_1, x_2) . The black circles denote the boundary of H_0 . The history structure inside H_0 between times $[-0.75, 1]$ is clearly visible in grey.

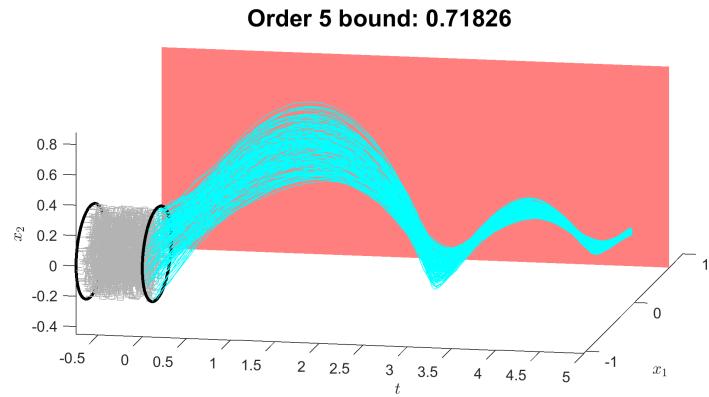


Figure 11.12: A 3d plot of (11.31) and its x_1 bound

The peak estimation of $p = x_2$ under the same system parameters leads to the sequence of

five bounds $p_{1:5}^* = [1.25, 1.25, 0.9557, 0.9138, 0.9112]$.

11.7 Conclusion

This chapter presented a formulation of MV-solutions for free-terminal-time DDEs with multiple histories (Definition 11.3.1). These MV-solutions are formed by the conjunction of Validity, Liouville and Consistency constraints. These MV-solutions may be used to provide upper bounds on peak estimation problems over DDEs by Program (11.11).

A vital area for future work is determining the conditions under which $P^* = p^*$ between (11.1) and (11.11) (Conjecture 11.3.1). Other areas for future work include applying MV-solutions to other problems (such as reachable set and positive-invariant set estimation) and formulating delay-dependent MV-solutions.

Chapter 12

Conclusion

This thesis solved peak estimation and safety quantification problems by extending an existing occupation measure framework. The measure LPs are successively approximated by a sequence of SDPs. Structure may be exploited (if present) to reduce the computational complexity of solving the SDPs. Combinations of peak estimation variations can be performed by merging together constraints in the measure LP formulations. Combinations of uncertainty and distance estimation may be particularly valuable in aeronautic applications, such as when there are uncertain wind patterns and the unsafe set is a mountain.

Part 1 reviewed the theory of peak estimation, introduced a rank-based recovery algorithm to attempt extraction of optimal trajectories, and provided an OCP framework for peak estimation with bounded uncertainties (time-independent, time-dependent, and switching). Part 1 concluded by adapting peak estimation for the safety margin and distance-of-closest-approach safety quantification schemes.

Part 2 formulated infinite-dimensional robust counterparts for Lie constraints, in order to eliminate the input-affine and SDR-constrained uncertainties and to render SOS tightenings of the LP computationally tractable. The robust counterparts have applications in data-driven peak and reachable set estimation. Peak-minimizing control was reviewed to define the crash-based safety quantification technique and the distance-maximizing control problem. Both crash-safety and distance-maximizing-control use the robust counterpart theory to generate solvable programs.

Part 3 applied peak estimation to non-ODE models. Hybrid systems were treated based on existing hybrid OCP methods, to which Zeno caps were introduced to ensure that the measure LP remained bounded. Peak estimation was adapted to SDEs by upper bounding the VaR using the chance-peak framework. DDEs were approached for peak estimation by defining free-terminal-time

CHAPTER 12. CONCLUSION

MV-solutions.

The main aspect for future work is in speeding up the execution time of SDPs derived from the moment-SOS hierarchy. This could include finding additional decomposable structure that could be used to bring down the runtime (e.g., network [64]), using Spectral Bundle methods [174, 175] rather than Interior Point Methods, Harmonic Hierarchies for polynomial optimization [176], or employing non-symmetric interior point methods [177]. One goal of such methods is to use the results of Chapter 11 to perform peak estimation of large-scale epidemic networks. Further effort is needed in Chapter 8 regarding experiments and implementation. Other research focuses include closing theoretical gaps, such as in performing a no-relaxation-gap proof for DDE peak estimation and optimal control (Conjecture 11.3.1), and in improving the VaR bounds for SDE using high-degree moments.

Bibliography

- [1] J. Miller and B. Shafai, “A Model of Heave Dynamics for Bagged Air Cushioned Vehicles,” in *2019 IEEE Conference on Control Technology and Applications (CCTA)*, 2019, pp. 976–981, [[link](#)].
- [2] J. Miller, M. A. Al-Radhawi, and E. D. Sontag, “Mediating Ribosomal Competition by Splitting Pools,” *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1555–1560, 2021.
- [3] J. Miller, Y. Zheng, B. Roig-Solvas, M. Sznaier, and A. Papachristodoulou, “Chordal Decomposition in Rank Minimized Semidefinite Programs with Applications to Subspace Clustering,” in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 4916–4921, [[link](#)].
- [4] J. Miller, Y. Zheng, M. Sznaier, and A. Papachristodoulou, “Decomposed structured subsets for semidefinite and sum-of-squares optimization,” *Automatica*, vol. 137, pp. 110–125, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005109821006543>
- [5] M. J. Cho and R. H. Stockbridge, “Linear Programming Formulation for Optimal Stopping Problems,” *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1965–1982, 2002.
- [6] G. Fantuzzi and D. Goluskin, “Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization,” *SIAM Journal on Applied Dynamical Systems*, vol. 19, no. 3, pp. 1823–1864, 2020.
- [7] R. Lewis and R. Vinter, “Relaxation of Optimal Control Problems to Equivalent Convex Programs,” *Journal of Mathematical Analysis and Applications*, vol. 74, no. 2, pp. 475–493, 1980.
- [8] D. Henrion, J. B. Lasserre, and C. Savorgnan, “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 4749–4754.

BIBLIOGRAPHY

- [9] G. Stewart, K. Heusden, and G. A. Dumont, “How Control Theory Can Help Us Control COVID-19,” *IEEE Spectrum*, vol. 57, no. 6, pp. 22–29, 2020.
- [10] D. H. Morris, F. W. Rossine, J. B. Plotkin, and S. A. Levin, “Optimal, near-optimal, and robust epidemic control,” *Communications Physics*, vol. 4, no. 1, pp. 1–8, 2021.
- [11] M. Sadeghi, J. M. Greene, and E. D. Sontag, “Universal features of epidemic models under social distancing guidelines,” *Annual reviews in control*, 2021.
- [12] E. Molina, A. Rapaport, and H. Ramírez, “Equivalent Formulations of Optimal Control Problems with Maximum Cost and Applications,” *Journal of Optimization Theory and Applications*, vol. 195, no. 3, pp. 953–975, 2022.
- [13] J. Miller and M. Sznaier, “Bounding the Distance to Unsafe Sets with Convex Optimization,” 2021, [[link](#)].
- [14] T. Tao, *An Introduction to Measure Theory*. American Mathematical Society Providence, RI, 2011, vol. 126.
- [15] M. Pivato, “Analysis, Measure, and Probability: A visual introduction,” 2003.
- [16] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, “Nonlinear Optimal Control via Occupation Measures and LMI-Relaxations,” *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [17] E. J. Anderson and P. Nash, *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley & Sons, 1987.
- [18] J. Cimprič, M. Marshall, and T. Netzer, “Closures of Quadratic Modules,” *Automatica*, vol. 183, no. 1, pp. 445–474, 2011.
- [19] M. Putinar, “Positive Polynomials on Compact Semi-algebraic Sets,” *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.
- [20] D. Henrion, M. Korda, and J. B. Lasserre, *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*. World Scientific, 2020, vol. 4.

BIBLIOGRAPHY

- [21] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” *IEEE Transactions on Automatic Control*, vol. 59, no. 2, p. 297–312, Feb 2014. [Online]. Available: <http://dx.doi.org/10.1109/TAC.2013.2283095>
- [22] M. Trnovská, “Strong duality conditions in semidefinite programming,” *Journal of Electrical Engineering*, vol. 56, no. 12, pp. 1–5, 2005.
- [23] M. Tacchi, “Convergence of Lasserre’s hierarchy: the general case,” *Optimization Letters*, vol. 16, no. 3, pp. 1015–1033, 2022.
- [24] R. E. Curto and L. A. Fialkow, *Solution of the truncated complex moment problem for flat data*. American Mathematical Soc., 1996, vol. 568.
- [25] ———, *Flat Extensions of Positive Moment Matrices: Recursively Generated Relations: Recursively Generated Relations*. American Mathematical Soc., 1998, vol. 648.
- [26] R. Curto and L. Fialkow, “The truncated complex \mathbb{K} -moment problem,” *Transactions of the American mathematical society*, vol. 352, no. 6, pp. 2825–2855, 2000.
- [27] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.
- [28] H. O. Fattorini, *Infinite Dimensional Optimization and Control Theory*. Cambridge University Press, 1999, vol. 54.
- [29] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*. SIAM, 2002.
- [30] D. Henrion and J.-B. Lasserre, “GloptiPoly: Global Optimization over Polynomials with MATLAB and SeDuMi,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 29, no. 2, pp. 165–194, 2003.
- [31] S. Marx, E. Pauwels, T. Weisser, D. Henrion, and J. Lasserre, “Tractable semi-algebraic approximation using christoffel-darboux kernel,” *arXiv preprint arXiv:1904.01833*, 2019.
- [32] J. Miller, D. Henrion, and M. Sznaier, “Peak Estimation Recovery and Safety Analysis,” *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 1982–1987, 2020.
- [33] D. Henrion and J.-B. Lasserre, “Detecting global optimality and extracting solutions in gloptipoly,” in *Positive Polynomials in Control*. Springer, Berlin, Heidelberg, 2005, vol. 312, pp. 293–310.

BIBLIOGRAPHY

- [34] H. García, C. Hernández, M. Junca, and M. Velasco, “Approximate super-resolution of positive measures in all dimensions,” *Applied and Computational Harmonic Analysis*, vol. 52, pp. 251–278, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1063520320300087>
- [35] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, “Exploiting Symmetries in SDP-Relaxations for Polynomial Optimization,” *Mathematics of Operations Research*, vol. 38, no. 1, p. 122–141, Feb 2013. [Online]. Available: <http://dx.doi.org/10.1287/moor.1120.0558>
- [36] S. Prajna and A. Jadbabaie, “Safety verification of hybrid systems using barrier certificates,” in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2004, pp. 477–492.
- [37] C. Schlosser and M. Korda, “Converging outer approximations to global attractors using semidefinite programming,” 2020.
- [38] J. C. Robinson, *Infinite-Dimensional Dynamical Systems: An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*. Cambridge University Press, 2001, vol. 28.
- [39] D. Goluskin, “Bounding extrema over global attractors using polynomial optimisation,” *Nonlinearity*, vol. 33, no. 9, p. 4878, 2020.
- [40] M. Tacchi, “Moment-sos hierarchy for large scale set approximation. application to power systems transient stability analysis,” Ph.D. dissertation, Toulouse, INSA, 2021.
- [41] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems*. Springer, 1996.
- [42] D. Henrion, M. Ganet-Schoeller, and S. Bennani, “Measures and LMI for space launcher robust control validation,” *IFAC Proceedings Volumes*, vol. 45, no. 13, pp. 236–241, 2012.
- [43] J. Miller, D. Henrion, M. Sznajer, and M. Korda, “Peak estimation for uncertain and switched systems,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 3222–3228.
- [44] K. Helmes, S. Röhl, and R. H. Stockbridge, “Computing Moments of the Exit Time Distribution for Markov Processes by Linear Programming,” *Operations Research*, vol. 49, no. 4, pp. 516–530, 2001.

BIBLIOGRAPHY

- [45] L. C. Young, “Generalized Surfaces in the Calculus of Variations,” *Annals of mathematics*, vol. 43, pp. 84–103, 1942.
- [46] P. A. Parrilo and A. Jadbabaie, “Approximation of the joint spectral radius using sum of squares,” *Linear Algebra and its Applications*, vol. 428, no. 10, pp. 2385–2402, 2008.
- [47] N. Vlassis and R. Jungers, “Polytopic uncertainty for linear systems: New and old complexity results,” *Systems & Control Letters*, vol. 67, pp. 9–13, 2014.
- [48] J. Löfberg, “YALMIP : A Toolbox for Modeling and Optimization in MATLAB,” in *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.
- [49] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.*, 2020. [Online]. Available: <https://docs.mosek.com/9.2/toolbox/index.html>
- [50] B. Xue, M. Fränzle, and N. Zhan, “Inner-Approximating Reachable Sets for Polynomial Systems with Time-Varying Uncertainties,” *IEEE Trans. Autom. Control*, vol. 65, no. 4, pp. 1468–1483, 2019.
- [51] A. Gorily, *Integrability and Nonintegrability of Dynamical Systems*. World Scientific, 2001, vol. 19.
- [52] V. Magron, P.-L. Garoche, D. Henrion, and X. Thirioux, “Semidefinite Approximations of Reachable Sets for Discrete-time Polynomial Systems,” *SIAM J. Control Optim.*, vol. 57, pp. 2799–2820, 03 2017.
- [53] C. Villani, *Optimal Transport: Old and New*. Springer Science & Business Media, 2008, vol. 338.
- [54] F. Santambrogio, “Optimal Transport for Applied Mathematicians,” *Birkhäuser, NY*, vol. 55, no. 58-63, p. 94, 2015.
- [55] G. Peyré, M. Cuturi *et al.*, “Computational Optimal Transport: With Applications to Data Science,” *Foundations and Trends® in Machine Learning*, vol. 11, no. 5-6, pp. 355–607, 2019.
- [56] S. Prajna, “Barrier certificates for nonlinear model validation,” *Automatica*, vol. 42, no. 1, pp. 117–126, 2006.

BIBLIOGRAPHY

- [57] A. Rantzer and S. Prajna, “On Analysis and Synthesis of Safe Control Laws,” in *42nd Allerton Conference on Communication, Control, and Computing*. University of Illinois, 2004, pp. 1468–1476.
- [58] J. Miller and M. Sznaier, “Bounding the Distance of Closest Approach to Unsafe Sets with Occupation Measures,” in *2022 61st IEEE Conference on Decision and Control (CDC)*, 2022, pp. 5008–5013.
- [59] M. M. Deza and E. Deza, “Encyclopedia of Distances,” in *Encyclopedia of distances*. Springer, Berlin, Heidelberg, 2009, pp. 1–583.
- [60] J. B. Lasserre, “A “Joint+Marginal” Approach to Parametric Polynomial Optimization,” *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1995–2022, 2010.
- [61] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity,” *SIOPT*, vol. 17, no. 1, pp. 218–242, 2006.
- [62] J. Wang, V. Magron, and J.-B. Lasserre, “TSSOS: A Moment-SOS hierarchy that exploits term sparsity,” *SIAM J. Optim.*, vol. 31, no. 1, pp. 30–58, 2021.
- [63] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, “Exploiting Symmetries in SDP-Relaxations for Polynomial Optimization,” *Mathematics of Operations Research*, vol. 38, no. 1, pp. 122–141, 2013.
- [64] C. Schlosser and M. Korda, “Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence,” 2020.
- [65] J. Wang, V. Magron, and J.-B. Lasserre, “Chordal-TSSOS: A Moment-SOS Hierarchy That Exploits Term Sparsity with Chordal Extension,” *SIAM J. Optim.*, vol. 31, no. 1, pp. 114–141, 2021.
- [66] L. Vandenberghe, M. S. Andersen *et al.*, “Chordal Graphs and Semidefinite Optimization,” *Foundations and Trends® in Optimization*, vol. 1, no. 4, pp. 241–433, 2015.
- [67] J. B. Lasserre, “Convergent SDP-Relaxations in Polynomial Optimization with Sparsity,” *SIAM Journal on Optimization*, vol. 17, no. 3, pp. 822–843, 2006.

BIBLIOGRAPHY

- [68] J. Lofberg, “YALMIP : a toolbox for modeling and optimization in MATLAB,” in *ICRA (IEEE Cat. No.04CH37508)*, 2004, pp. 284–289.
- [69] J. Miller and M. Sznajer, “Facial Input Decompositions for Robust Peak Estimation under Polyhedral Uncertainty,” *IFAC-PapersOnLine*, vol. 55, no. 25, pp. 55–60, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896322015750>
- [70] M. Yannakakis, “Expressing combinatorial optimization problems by Linear Programs,” *Journal of Computer and System Sciences*, vol. 43, no. 3, pp. 441–466, 1991.
- [71] J. Gouveia, P. A. Parrilo, and R. R. Thomas, “Lifts of Convex Sets and Cone Factorizations,” *Mathematics of Operations Research*, vol. 38, no. 2, pp. 248–264, 2013.
- [72] D. Anderson and M. Osborne, “Discrete, linear approximation problems in polyhedral norms,” *Numerische Mathematik*, vol. 26, no. 2, pp. 179–189, 1976.
- [73] J. W. Helton and J. Nie, “Semidefinite representation of convex sets,” *Mathematical Programming*, vol. 122, pp. 21–64, 2010.
- [74] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009, vol. 28.
- [75] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.
- [76] A. Ben-Tal, D. Den Hertog, and J.-P. Vial, “Deriving Robust Counterparts of Nonlinear Uncertain Inequalities,” *Mathematical programming*, vol. 149, no. 1-2, pp. 265–299, 2015.
- [77] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake, “Convex Optimization of Nonlinear Feedback Controllers via Occupation Measures,” *The International Journal of Robotics Research*, vol. 33, no. 9, pp. 1209–1230, 2014.
- [78] M. Korda, D. Henrion, and C. N. Jones, “Controller design and value function approximation for nonlinear dynamical systems,” *Automatica*, vol. 67, pp. 54–66, 2016.
- [79] J. Miller and M. Sznajer, “Analysis and Control of Input-Affine Dynamical Systems using Infinite-Dimensional Robust Counterparts,” 2023, arxiv:2112.14838.
- [80] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory: new directions*. Springer Science & Business Media, 2011.

BIBLIOGRAPHY

- [81] M. Korda, D. Henrion, and C. N. Jones, “Inner approximations of the region of attraction for polynomial dynamical systems,” *IFAC Proceedings Volumes*, vol. 46, no. 23, pp. 534–539, 2013.
- [82] A. Majumdar, A. A. Ahmadi, and R. Tedrake, “Control Design along Trajectories with Sums of Squares Programming,” in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 4054–4061.
- [83] A. B. Kurzhanski and P. Varaiya, “On ellipsoidal techniques for reachability analysis. part i: external approximations,” *Optimization methods and software*, vol. 17, no. 2, pp. 177–206, 2002.
- [84] S. M. Harwood and P. I. Barton, “Efficient polyhedral enclosures for the reachable set of nonlinear control systems,” *Mathematics of Control, Signals, and Systems*, vol. 28, no. 1, p. 8, 2016.
- [85] S. Coogan, “Mixed monotonicity for reachability and safety in dynamical systems,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 5074–5085.
- [86] T. Dai and M. Sznaier, “A Semi-Algebraic Optimization Approach to Data-Driven Control of Continuous-Time Nonlinear Systems,” *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 487–492, 2020.
- [87] C. W. Scherer and C. W. Hol, “Matrix Sum-of-Squares Relaxations for Robust Semi-Definite Programs,” *Mathematical programming*, vol. 107, no. 1, pp. 189–211, 2006.
- [88] C. Hol and C. Scherer, “Sum of squares relaxations for robust polynomial semi-definite programs,” *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 451–456, 2005.
- [89] F. Alizadeh and D. Goldfarb, “Second-order cone programming,” *Mathematical programming*, vol. 95, no. 1, pp. 3–51, 2003.
- [90] Y. Cheng, M. Sznaier, and C. Lagoa, “Robust Superstabilizing Controller Design from Open-Loop Experimental Input/Output Data,” *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 1337–1342, 2015.
- [91] T. Dai and M. Sznaier, “A Moments Based Approach to Designing MIMO Data Driven Controllers for Switched Systems,” in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 5652–5657.

BIBLIOGRAPHY

- [92] T. Martin, T. B. Schön, and F. Allgöwer, “Gaussian inference for data-driven state-feedback design of nonlinear systems,” *arXiv preprint arXiv:2211.05639*, 2022.
- [93] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, “From noisy data to feedback controllers: non-conservative design via a Matrix S-Lemma,” *IEEE Trans. Automat. Contr.*, 2020.
- [94] R. Caron, J. McDonald, and C. Ponic, “A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary,” *Journal of Optimization Theory and Applications*, vol. 62, no. 2, pp. 225–237, 1989.
- [95] D. P. Kroese, T. Taimre, and Z. I. Botev, *Handbook of monte carlo methods*. John Wiley & Sons, 2013, vol. 706.
- [96] T. Benham, “Uniform distribution over a convex polytope,” 2012.
[Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/34208-uniform-distribution-over-a-convex-polytope>
- [97] M. Korda, D. Henrion, and C. Jones, “Convex Computation of the Maximum Controlled Invariant Set For Polynomial Control Systems,” *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2944–2969, 2014.
- [98] Z. Qu and X. Tang, “A correlative sparse lagrange multiplier expression relaxation for polynomial optimization,” *arXiv preprint arXiv:2208.03979*, 2022.
- [99] R. B. Vinter, “Minimax optimal control,” *SIAM journal on control and optimization*, vol. 44, no. 3, pp. 939–968, 2005.
- [100] E. Molina and A. Rapaport, “An optimal feedback control that minimizes the epidemic peak in the sir model under a budget constraint,” *Automatica*, vol. 146, p. 110596, 2022.
- [101] P. Lu and N. X. Vinh, “Minimax optimal control for atmospheric fly-through trajectories,” *Journal of Optimization Theory and Applications*, vol. 57, no. 1, pp. 41–58, 1988.
- [102] H. Kreim, B. Kugelmann, H. J. Pesch, and M. H. Breitner, “Minimizing the Maximum Heating of a Reentering Space Shuttle: An Optimal Control Problem with Multiple Control Constraints,” *Optimal Control Applications and Methods*, vol. 17, no. 1, pp. 45–69, 1996.
- [103] J. Miller and M. Sznaier, “Quantifying the Safety of Trajectories using Peak-Minimizing Control,” 2023, arxiv:2303.11896.

BIBLIOGRAPHY

- [104] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, “CasADi: a software framework for nonlinear optimization and optimal control,” *Mathematical Programming Computation*, vol. 11, pp. 1–36, 2019.
- [105] R. B. Vinter and R. M. Lewis, “A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumptions on the controls,” *SIAM Journal on Control and Optimization*, vol. 16, no. 4, pp. 571–583, 1978.
- [106] M. Jones and M. M. Peet, “Polynomial Approximation of Value Functions and Nonlinear Controller Design with Performance Bounds,” 2021.
- [107] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2011.
- [108] D. Gottlieb and C.-W. Shu, “On the Gibbs phenomenon and its resolution,” *SIAM review*, vol. 39, no. 4, pp. 644–668, 1997.
- [109] M. Tacchi, J. B. Lasserre, and D. Henrion, “Stokes, gibbs, and volume computation of semi-algebraic sets,” *Discrete & Computational Geometry*, pp. 1–24, 2022.
- [110] K. Gatermann and P. A. Parrilo, “Symmetry groups, semidefinite programs, and sums of squares,” *Journal of Pure and Applied Algebra*, vol. 192, no. 1-3, pp. 95–128, 2004.
- [111] R. Goebel, R. G. Sanfelice, and A. R. Teel, “Hybrid dynamical systems,” *IEEE control systems magazine*, vol. 29, no. 2, pp. 28–93, 2009.
- [112] S. Collins, A. Ruina, R. Tedrake, and M. Wisse, “Efficient bipedal robots based on passive-dynamic walkers,” *Science*, vol. 307, no. 5712, pp. 1082–1085, 2005.
- [113] N. Zaupa, L. Martínez-Salamero, C. Olalla, and L. Zaccarian, “Results on hybrid control of self-oscillating resonant converters,” *IFAC-PapersOnLine*, vol. 54, no. 5, pp. 211–216, 2021, 7th IFAC Conference on Analysis and Design of Hybrid Systems ADHS 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896321012751>
- [114] P. Naghshtabrizi, J. P. Hespanha, and A. R. Teel, “On the robust stability and stabilization of sampled-data systems: A hybrid system approach,” in *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006, pp. 4873–4878.

BIBLIOGRAPHY

- [115] S. Fischer, P. Kurbatova, N. Bessonov, O. Gandrillon, V. Volpert, and F. Crauste, “Modeling erythroblastic islands: Using a hybrid model to assess the function of central macrophage,” *Journal of Theoretical Biology*, vol. 298, pp. 92–106, 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022519312000033>
- [116] P. Zhao, S. Mohan, and R. Vasudevan, “Optimal control of polynomial hybrid systems via convex relaxations,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2062–2077, 2019.
- [117] P. Zhao and R. Vasudevan, “Nonlinear hybrid optimal control with switching costs via occupation measures and lmi-relaxations,” in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 4293–4300.
- [118] A. Rocca, M. Forets, V. Magron, E. Fanchon, and T. Dang, “Occupation measure methods for modelling and analysis of biological hybrid systems,” *IFAC-PapersOnLine*, vol. 51, no. 16, pp. 181–186, 2018, 6th IFAC Conference on Analysis and Design of Hybrid Systems ADHS 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896318311467>
- [119] V. Shia, R. Vasudevan, R. Bajcsy, and R. Tedrake, “Convex computation of the reachable set for controlled polynomial hybrid systems,” in *53rd IEEE Conference on Decision and Control*, 2014, pp. 1499–1506.
- [120] S. Mohan, V. Shia, and R. Vasudevan, “Convex computation of the reachable set for hybrid systems with parametric uncertainty,” 2016.
- [121] J. Miller and M. Sznajer, “Peak Estimation of Hybrid Systems with Convex Optimization,” 2023, arxiv:2303.11490.
- [122] N. Hartsfield and G. Ringel, *Pearls in graph theory: a comprehensive introduction*. Courier Corporation, 2013.
- [123] A. D. Ames, H. Zheng, R. D. Gregg, and S. Sastry, “Is there life after zeno? taking executions past the breaking (zeno) point,” in *2006 American control conference*. IEEE, 2006, pp. 6–pp.
- [124] W. Han and R. Tedrake, “Controller synthesis for discrete-time hybrid polynomial systems via occupation measures,” in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 7675–7682.

BIBLIOGRAPHY

- [125] P. Jorion, *Value at Risk*. McGraw-Hill Professional Publishing, 2000.
- [126] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, “Coherent measures of risk,” *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.
- [127] G. C. Pflug, “Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk,” in *Probabilistic constrained optimization*. Springer, 2000, pp. 272–281.
- [128] R. T. Rockafellar and S. Uryasev, “Conditional Value-at-Risk for General Loss Distributions,” *Journal of banking & finance*, vol. 26, no. 7, pp. 1443–1471, 2002.
- [129] C. W. Miller and I. Yang, “Optimal control of conditional value-at-risk in continuous time,” *SIAM Journal on Control and Optimization*, vol. 55, no. 2, pp. 856–884, 2017.
- [130] M. P. Chapman, R. Bonalli, K. M. Smith, I. Yang, M. Pavone, and C. J. Tomlin, “Risk-sensitive safety analysis using conditional value-at-risk,” *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6521–6536, 2021.
- [131] A. Charnes, W. W. Cooper, and G. H. Symonds, “Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil,” *Management science*, vol. 4, no. 3, pp. 235–263, 1958.
- [132] C. M. Lagoa, X. Li, and M. Sznaier, “Probabilistically Constrained Linear Programs and Risk-Adjusted Controller Design,” *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 938–951, 2005.
- [133] F. P. Cantelli, “Sui confini della probabilità,” in *Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928*, 1929, pp. 47–60.
- [134] D. Vysotskij and Y. I. Petunin, “Justification of the 3σ rule for unimodal distributions,” *Theory of Probability and Mathematical Statistics*, vol. 21, no. 25-36, 1980.
- [135] A. Wang, A. Jasour, and B. C. Williams, “Non-gaussian chance-constrained trajectory planning for autonomous vehicles under agent uncertainty,” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6041–6048, 2020.
- [136] W. Han, A. Jasour, and B. Williams, “Non-Gaussian Risk Bounded Trajectory Optimization for Stochastic Nonlinear Systems in Uncertain Environments,” *arXiv preprint arXiv:2203.03038*, 2022.

BIBLIOGRAPHY

- [137] G. Calafiore and M. C. Campi, “Uncertain convex programs: randomized solutions and confidence levels,” *Mathematical Programming*, vol. 102, no. 1, pp. 25–46, 2005.
- [138] A. M. Jasour, N. S. Aybat, and C. M. Lagoa, “Semidefinite Programming For Chance Constrained Optimization Over Semialgebraic Sets,” *SIAM Journal on Optimization*, vol. 25, no. 3, pp. 1411–1440, 2015.
- [139] K. Kashima and R. Kawai, “An optimization approach to weak approximation of Lévy-driven stochastic differential equations,” in *Perspectives in Mathematical System Theory, Control, and Signal Processing*. Springer, 2010, pp. 263–272.
- [140] J.-B. Lasserre, T. Prieto-Rumeau, and M. Zervos, “Pricing a class of exotic options via moments and SDP relaxations,” *Mathematical Finance*, vol. 16, no. 3, pp. 469–494, 2006.
- [141] S. Prajna, A. Jadbabaie, and G. J. Pappas, “A Framework for Worst-Case and Stochastic Safety Verification Using Barrier Certificates,” *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415–1428, 2007.
- [142] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko, “Bounds for Deterministic and Stochastic Dynamical Systems using Sum-of-Squares Optimization,” *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 4, pp. 1962–1988, 2016.
- [143] B. Xue, N. Zhan, and M. Fränzle, “Reach-Avoid Analysis for Stochastic Differential Equations,” 2022.
- [144] J. Miller, M. Tacchi, M. Sznaier, and A. Jasour, “Peak Value-at-Risk Estimation for Stochastic Differential Equations using Occupation Measures,” 2023, arxiv:2303.16064.
- [145] C. Gardiner, *Stochastic methods*. Springer Berlin, 2009, vol. 4.
- [146] B. Øksendal, “Stochastic Differential Equations: An Introduction with Applications ,” in *Stochastic differential equations*. Springer, 2003, pp. 65–84.
- [147] R. Fürth, “Einige Untersuchungen über Brownsche Bewegung an einem Einzelteilchen,” *Annalen der Physik*, vol. 358, no. 11, pp. 177–213, 1917.
- [148] H. J. Hwang, J. Jang, and J. J. Velázquez, “The Fokker–Planck equation with absorbing boundary conditions,” *Archive for Rational Mechanics and Analysis*, vol. 214, no. 1, pp. 183–233, 2014.

BIBLIOGRAPHY

- [149] E. B. Dynkin, “Markov Processes,” in *Markov processes*. Springer, 1965, pp. 77–104.
- [150] J. Lofberg. (2009) Working with square roots. [Online]. Available: <https://yalmip.github.io/squareroots>
- [151] D. Henrion, M. Junca, and M. Velasco, “Moment-SOS hierarchy and exit time of stochastic processes,” *arXiv preprint arXiv:2101.06009*, 2021.
- [152] S. Prajna, A. Jadbabaie, and G. J. Pappas, “Stochastic safety verification using barrier certificates,” in *2004 43rd IEEE conference on decision and control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 1. IEEE, 2004, pp. 929–934.
- [153] S. He, J. Zhang, and S. Zhang, “Bounding probability of small deviation: A fourth moment approach,” *Mathematics of Operations Research*, vol. 35, no. 1, pp. 208–232, 2010.
- [154] D. Applebaum, “Lévy processes from probability to finance and quantum groups,” *Notices of the AMS*, vol. 51, no. 11, pp. 1336–1347, 2004.
- [155] J. K. Hale, “Functional Differential Equations,” in *Analytic theory of differential equations*. Springer, 1971, pp. 9–22.
- [156] Y. Kuang, *Delay Differential Equations: with Applications in Population Dynamics*. Academic press, 1993.
- [157] A. Bellen and M. Zennaro, *Numerical Methods for Delay Differential Equations*. Oxford university press, 2013.
- [158] E. Fridman, *Introduction to Time-Delay Systems: Analysis and Control*. Birkhäuser Basel, 10 2014.
- [159] J. Warga, “Optimal Controls with Pseudodelays,” *SIAM Journal on Control*, vol. 12, no. 2, pp. 286–299, 1974.
- [160] ——, *Optimal control of differential and functional equations*. Academic press, 2014.
- [161] J. F. Rosenblueth and R. B. Vinter, “Relaxation Procedures for Time Delay Systems,” *Journal of Mathematical Analysis and Applications*, vol. 162, no. 2, pp. 542–563, 1991.
- [162] J. F. Rosenblueth, “Strongly and weakly relaxed controls for time delay systems,” *SIAM Journal on Control and Optimization*, vol. 30, no. 4, pp. 856–866, 1992.

BIBLIOGRAPHY

- [163] J. Rosenblueth, “Proper relaxation of optimal control problems,” *Journal of Optimization Theory and Applications*, vol. 74, no. 3, pp. 509–526, 1992.
- [164] A. Papachristodoulou and S. Prajna, “A Tutorial on Sum of Squares Techniques for Systems Analysis,” in *Proceedings of the 2005, American Control Conference, 2005*. IEEE, 2005, pp. 2686–2700.
- [165] S. Prajna and A. Jadbabaie, “Methods for Safety Verification of Time-Delay Systems,” in *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 4348–4353.
- [166] A. Papachristodoulou, M. M. Peet, and S. Lall, “Analysis of Polynomial Systems With Time Delays via the Sum of Squares Decomposition,” *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1058–1064, 2009.
- [167] S. Marx, T. Weisser, D. Henrion, and J. B. Lasserre, “A moment approach for entropy solutions to nonlinear hyperbolic PDEs,” *Mathematical Control and Related Fields*, vol. 10, no. 1, pp. 113–140, 2020.
- [168] M. Korda, D. Henrion, and J. B. Lasserre, “Chapter 10 - Moments and convex optimization for analysis and control of nonlinear PDEs,” in *Numerical Control: Part A*, ser. Handbook of Numerical Analysis, E. Trélat and E. Zuazua, Eds. Elsevier, 2022, vol. 23, pp. 339–366. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1570865921000259>
- [169] V. Magron and C. Prieur, “Optimal control of linear PDEs using occupation measures and SDP relaxations,” *IMA Journal of Mathematical Control and Information*, vol. 37, no. 1, pp. 159–174, 2020.
- [170] S. Barati, “Optimal Control of Constrained Time Delay Systems,” *Advanced Modeling and Optimization*, vol. 14, no. 1, pp. 103–116, 2012.
- [171] R. Suresh and V. K. Chandrasekar, “Influence of time-delay feedback on extreme events in a forced Liénard system,” *Phys. Rev. E*, vol. 98, p. 052211, Nov 2018.
- [172] J. Miller, M. Korda, V. Magron, and M. Sznaier, “Peak Estimation of Time Delay Systems using Occupation Measures,” 2023, arxiv:2303.12863. [Online]. Available: <https://arxiv.org/abs/2303.12863>

BIBLIOGRAPHY

- [173] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, “The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application,” *Annals of internal medicine*, vol. 172, no. 9, pp. 577–582, 2020.
- [174] N. H. A. Mai, J.-B. Lasserre, V. Magron, and J. Wang, “Exploiting constant trace property in large-scale polynomial optimization,” *ACM Transactions on Mathematical Software*, vol. 48, no. 4, pp. 1–39, 2022.
- [175] A. Yurtsever, J. A. Tropp, O. Fercoq, M. Udell, and V. Cevher, “Scalable semidefinite programming,” *SIAM Journal on Mathematics of Data Science*, vol. 3, no. 1, pp. 171–200, 2021.
- [176] S. Cristancho and M. Velasco, “Harmonic hierarchies for polynomial optimization,” *arXiv preprint arXiv:2202.12865*, 2022.
- [177] D. Papp and S. Yildiz, “Sum-of-squares optimization without semidefinite programming,” *SIAM Journal on Optimization*, vol. 29, no. 1, pp. 822–851, 2019.
- [178] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*. Springer Science & Business Media, 2009.
- [179] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer Science & Business Media, 2009, vol. 317.
- [180] J. Denel, “Extensions of the continuity of point-to-set maps: applications to fixed point algorithms,” *Point-to-Set Maps and Mathematical Programming*, pp. 48–68, 1979.
- [181] O. L. Mangasarian and T.-H. Shiao, “Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems,” *SIAM Journal on Control and Optimization*, vol. 25, no. 3, pp. 583–595, 1987.
- [182] J. G. Llavona, *Approximation of continuously differentiable functions*. Elsevier, 1986.
- [183] R. Henrion and A. Seeger, “On properties of different notions of centers for convex cones,” *Set-Valued and Variational Analysis*, vol. 18, pp. 205–231, 2010.
- [184] ——, “Inradius and circumradius of various convex cones arising in applications,” *Set-Valued and Variational Analysis*, vol. 18, pp. 483–511, 2010.

BIBLIOGRAPHY

- [185] A. Barvinok, *A Course in Convexity*. American Mathematical Society, 2002.
- [186] J. R. Ockendon and A. B. Tayler, “The dynamics of a current collection system for an electric locomotive,” *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 322, no. 1551, pp. 447–468, 1971.
- [187] J. Carr and J. Dyson, “2.-the matrix functional differential equation $y'(x) = ay(\lambda x) + b(x)$,” *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, vol. 75, no. 1, pp. 5–22, 1976.
- [188] A. Iserles, *On the generalized pantograph functional-differential equation*. University of Cambridge, Department of Applied Mathematics and Theoretical, 1991.
- [189] ——, “On nonlinear delay differential equations,” *Transactions of the American Mathematical Society*, vol. 344, no. 1, pp. 441–477, 1994.
- [190] Y. Liu, “Asymptotic behaviour of functional-differential equations with proportional time delays,” *European Journal of Applied Mathematics*, vol. 7, no. 1, pp. 11–30, 1996.
- [191] L. Shampine, “Solving odes and ddes with residual control,” *Applied Numerical Mathematics*, vol. 52, no. 1, pp. 113–127, 2005.
- [192] I. Ali, H. Brunner, and T. Tang, “A spectral method for pantograph-type delay differential equations and its convergence analysis,” *Journal of Computational Mathematics*, pp. 254–265, 2009.
- [193] S. Sedaghat, Y. Ordokhani, and M. Dehghan, “Numerical solution of the delay differential equations of pantograph type via chebyshev polynomials,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4815–4830, 2012.
- [194] M. M. Bahşı and M. Çevik, “Numerical solution of pantograph-type delay differential equations using perturbation-iteration algorithms,” *Journal of Applied Mathematics*, 2015.
- [195] A. Plaksin, “Minimax and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations for Time-Delay Systems,” *Journal of Optimization Theory and Applications*, vol. 187, no. 1, pp. 22–42, 2020.
- [196] O. Santos, S. Mondié, and V. L. Kharitonov, “Linear quadratic suboptimal control for time delays systems,” *International Journal of Control*, vol. 82, no. 1, pp. 147–154, 2009.

BIBLIOGRAPHY

- [197] J.-M. Ortega-Martínez, O.-J. Santos-Sánchez, and S. Mondié, “Comments on the bellman functional for linear time-delay systems,” *Optimal Control Applications and Methods*, vol. 42, no. 5, pp. 1531–1540, 2021.

Chapter A

Appendices

A.1 Proof of Strong Duality for Distance Estimation in Theorem 5.4.5

This proof will follow the method used in Theorem 2.6 of [40] to prove duality.

The two programs (5.11) and (5.17) will be posed as a pair of standard-form infinite-dimensional LPs using notation from [40]. The following spaces may be defined:

$$\begin{aligned}\mathcal{X}' &= C(X_0) \times C([0, T] \times X)^2 \times C(X \times X_u) \\ \mathcal{X} &= \mathcal{M}(X_0) \times \mathcal{M}([0, T] \times X)^2 \times \mathcal{M}(X \times X_u).\end{aligned}\tag{A.1}$$

The nonnegative subcones of \mathcal{X}' and \mathcal{X} respectively are

$$\begin{aligned}\mathcal{X}'_+ &= C_+(X_0) \times C_+([0, T] \times X)^2 \times C_+(X \times X_u) \\ \mathcal{X}_+ &= \mathcal{M}_+(X_0) \times \mathcal{M}_+([0, T] \times X)^2 \times \mathcal{M}_+(X \times X_u).\end{aligned}\tag{A.2}$$

The cones \mathcal{X}'_+ and \mathcal{X}_+ in (A.2) are topological duals under assumption A1, and the measures from (5.11e)-(5.11f) satisfy $\mu = (\mu_0, \mu_p, \mu, \eta) \in \mathcal{X}_+$. The spaces \mathcal{Y} and \mathcal{Y}' may be defined as

$$\mathcal{Y}' = C(X) \times C^1([0, T] \times X) \times \mathbb{R}\tag{A.3}$$

$$\mathcal{Y} = \mathcal{M}(X) \times C^1([0, T] \times X)' \times 0.\tag{A.4}$$

We express $\mathcal{Y}_+ = \mathcal{Y}$ and $\mathcal{Y}'_+ = \mathcal{Y}'$ to maintain a convention with [40] given there are no affine-inequality constraints in (5.11). We equip \mathcal{X} with the weak-* topology and \mathcal{Y} with the (sup-norm bounded) weak topology. The arguments $\ell = (w, v, \gamma)$ from problem (5.17) are members of the set \mathcal{Y}'_+ .

CHAPTER A. APPENDICES

The linear operators $\mathcal{A}' : \mathcal{Y}'_+ \rightarrow \mathcal{X}'_+$ and $\mathcal{A} : \mathcal{X}_+ \rightarrow \mathcal{Y}_+$ induced from constraints (5.11b)-(5.11d) may be defined as

$$\begin{aligned}\mathcal{A}(\boldsymbol{\mu}) &= [\pi_\#^x \mu_p - \pi_\#^x \eta, \delta_0 \otimes \mu_0 + \mathcal{L}_f^\dagger \mu - \mu_p, \langle 1, \mu_0 \rangle] \\ \mathcal{A}'(\boldsymbol{\ell}) &= [v(0, x) - \gamma, w(x) - v(t, x), \mathcal{L}_f v(t, x), -w(x)].\end{aligned}\quad (\text{A.5})$$

The last pieces needed to convert (5.11) into a standard-form LP are the cost vector $\mathbf{c} = [0, 0, 0, c(x, y)]$ and the answer vector $\mathbf{b} = [0, 0, 1] \in \mathcal{Y}'$. Problem (5.11) is therefore equivalent to (with $\langle \mathbf{c}, \boldsymbol{\mu} \rangle = \langle c, \eta \rangle$)

$$p^* = \inf_{\boldsymbol{\mu} \in \mathcal{X}_+} \langle \mathbf{c}, \boldsymbol{\mu} \rangle \quad \mathbf{b} - \mathcal{A}(\boldsymbol{\mu}) \in \mathcal{Y}_+. \quad (\text{A.6})$$

The dual LP to (A.6) in standard form is (with $\langle \boldsymbol{\ell}, \mathbf{b} \rangle = \gamma$)

$$d^* = \sup_{\boldsymbol{\ell} \in \mathcal{Y}'_+} \langle \boldsymbol{\ell}, \mathbf{b} \rangle \quad \mathcal{A}'(\boldsymbol{\ell}) - \mathbf{c} \in \mathcal{X}_+. \quad (\text{A.7})$$

The operators \mathcal{A} and \mathcal{A}' are adjoints with $\langle \mathcal{A}(\boldsymbol{\ell}), \boldsymbol{\mu} \rangle = \langle \boldsymbol{\ell}, \mathcal{A}'(\boldsymbol{\mu}) \rangle$ for all $\boldsymbol{\ell} \in \mathcal{Y}'_+$ and $\boldsymbol{\mu} \in \mathcal{X}_+$.

The sufficient conditions for strong duality and attainment of optimality between (A.6) and (A.7) as outlined in Theorem 2.6 of [40] are that:

R1 All support sets are compact (A1).

R2 All measure solutions have bounded mass (Lemma 5.5.1).

R3 All functions involved in the definitions of c and \mathcal{A} are continuous (A2, A3).

R4 There exists a $\boldsymbol{\mu}_{\text{feas}} \in \mathcal{X}_+$ with $\mathbf{b} - \mathcal{A}(\boldsymbol{\mu}_{\text{feas}}) \in \mathcal{Y}_+$.

The requirements R1 and R2 hold by Assumption A1 and Lemma 5.5.1 respectively. R3 is valid given that $c(x, y)$ is C^0 (A3), the projection map π^x is continuous, and the mapping $(t, x) \mapsto \mathcal{L}_f v(t, x)$ is C^0 for $v \in C^1$ and f Lipschitz (continuous) (A2). A feasible measure $\boldsymbol{\mu}_{\text{feas}}$ may be constructed from the process in Theorem 5.4.1 from a tuple \mathcal{T} , therefore satisfying R4.

Strong duality between (5.11) and (5.17) is therefore proven after satisfaction of all four requirements.

A.2 Continuity of Multipliers

This section will prove Theorem 6.4.3. Along the way, it will form a general condition for lower semicontinuity of robust counterparts.

CHAPTER A. APPENDICES

A.2.1 Set-Valued Preliminaries

We first review concepts in set-valued analysis. Given spaces Y and Z , a set-valued function $F : Y \rightrightarrows Z$ is a mapping between the power sets $F : 2^Y \rightarrow 2^Z$. A set $E \subset Y$ is inside the domain $\text{Dom}(F)$ if $F(E) \neq \emptyset$. In this section, we will be utilizing point-set maps ($F : Y \rightarrow 2^Z$).

Definition A.2.1 (Definition 1.4.2 of [178]). *The function F is **lower semicontinuous** at $y \in \text{Dom}(F)$ if, for every sequence $\{y_k\}$ converging to y ($\{y_k\} \rightarrow y$), there exists a converging sequence $\{z_k \in F(y)\}$ converging to an element $z \in F(y)$. The map F is lower semicontinuous if it is lower semicontinuous at each $y \in \text{Dom}(F)$.*

Definition A.2.2. *Let F_0, F_1 be set-valued maps $Y \rightrightarrows Z$. The containment relation $F_0 \subseteq F_1$ holds if $\forall y \in Y : F_0(y) \subseteq F_1(y)$.*

Remark A.2.1. *Lower semicontinuity in Definition A.2.1 is also called **inner semicontinuity** in [179].*

Definition A.2.3 (Definition 1 of [180]). *A family of set-valued maps $\{S_\rho : Y \rightrightarrows Z\}_{\rho \geq 0}$ is a **ρ -decreasing family** if $\forall \rho \geq \rho' \geq 0, y \in Y : S_\rho \subseteq S_{\rho'}$.*

Definition A.2.4 (Definition 2 of [180]). *A ρ -decreasing family $\{S_\rho\}_{\rho \geq 0}$ is **dense** if $S_0(y) \subseteq \text{Closure}(\cup_{\rho > 0} S_\rho(y))$ for all $y \in Y$.*

Definition A.2.5 (Definition 3 of [180]). *A ρ -decreasing family $\{S_\rho\}_{\rho \geq 0}$ is **pseudo-lower-continuous** at y if for all sequences $\{y_k\} \rightarrow y$, parameters $\rho > \rho' > 0$, and points $z \in S_\rho(y)$, there exists an $N \in \mathbb{N}$ and a sequence $\{z_k\} \rightarrow z$ such that $\forall k \geq N : z_k \in S_{\rho'}(y_k)$. The family is **pseudo-lower-continuous** if it is pseudo-lower-continuous at all $y \in Y$.*

Corollary 4 (Excerpt of Remark 5 of [180]). *If a ρ -decreasing family $\{S_\rho\}_{\rho \geq 0}$ is pseudo-lower-continuous and dense, then S_0 is lower semicontinuous.*

A.2.2 Lower Semicontinuity of Strict Robust Counterpart Multipliers

This subsection will analyze continuity properties of the strict semi-infinite inequality (6.7). The following assumptions are required:

- A1' The assumptions of Theorem 6.2.1 are satisfied (convex pointed cones with nonempty interiors satisfying Slater conditions).

CHAPTER A. APPENDICES

A2' The parameter set Y is compact.

A3' The problem data (a_0, b_0, b_ℓ, e_s) of (6.8) are all continuous functions of $y \in Y$.

A4' The problem data (a_ℓ, A_s, G_s) are all constant in y .

Remark A.2.2. *Continuity of problem data in A4' over Y implies that all problem entries are finite.*

Define the following quantities based on (6.6):

$$Z = K^* \times \mathbb{R}^L \quad e = [e_1; e_2; \dots; e_{N_s}] \quad (\text{A.8a})$$

$$a_\bullet = [a_1, a_2, \dots, a_L] \quad b_\bullet = [b_1; b_2; \dots; b_L] \quad (\text{A.8b})$$

$$A = \text{blkdiag}(A_1, A_2, \dots, A_{N_s}) \quad G = \text{blkdiag}(G_1, G_2, \dots, G_{N_s}). \quad (\text{A.8c})$$

We define a ρ -indexed family of set-valued maps $S_\rho : Y \rightrightarrows Z$ as the ρ -modified solution map to (6.6):

$$S_{\rho>0}(y) = \left\{ (\zeta, \beta) \in Z : \begin{array}{l} e^T \zeta + a_0^T \beta + \rho \leq b_0 \\ G^T \zeta = 0 \\ A^T \zeta + a_\bullet^T \beta = b_\bullet \end{array} \right\}, \quad (\text{A.9a})$$

$$S_0(y) = \left\{ (\zeta, \beta) \in Z : \begin{array}{l} e^T \zeta + a_0^T \beta < b_0 \\ G^T \zeta = 0 \\ A^T \zeta + a_\bullet^T \beta = b_\bullet \end{array} \right\}. \quad (\text{A.9b})$$

The semi-infinite strict program (6.7) has a solution at the parameter y if $S_0(y) \neq \emptyset$.

Lemma A.2.1. *The family $\{S_\rho\}_{\rho \geq 0}$ from (A.9) is a ρ -decreasing family (Def. A.2.3).*

Proof. The tolerance ρ only appears in the linear inequality $e^T \zeta + a_0^T \beta + \rho \leq b_0$. The maps therefore satisfy $\forall \rho > \rho' \geq 0 : S_\rho \subseteq S_{\rho'}$. \square

Theorem A.2.2. *The mapping S_0 from (A.9b) is lower semicontinuous.*

Proof. This theorem will be proved using Proposition 3 of [180]. Define the following maps:

$$\phi(y) = -b_0 \quad (\text{A.10a})$$

$$\gamma(\rho) = \rho \quad (\text{A.10b})$$

$$\xi(y, \zeta, \beta) = -(e^T \zeta + a_0^T \beta) \quad (\text{A.10c})$$

$$\Psi_\rho(y) = \left\{ (\zeta, \beta) \in Z : \begin{bmatrix} G^T & \mathbf{0} \\ A^T & a_\bullet^T \end{bmatrix} \begin{bmatrix} \zeta \\ \beta \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ b_\bullet \end{bmatrix} \right\}. \quad (\text{A.10d})$$

CHAPTER A. APPENDICES

Note that $\Psi_\rho(y)$ in (A.10d) is independent of ρ , that b_\bullet is a continuous function of y (A3'), and (G, A, a_\bullet) are constant in terms of y (A4'). The map $\Psi_\rho(y)$ is additionally lower-semicontinuous (and Lipschitz) w.r.t. perturbations in $b_\bullet(y)$ by extension of arguments in Theorem 2.2 of [181] to the conic case (noting that Ψ_ρ is constant in its left-hand side). The mapping S_ρ from (A.9) may be expressed in terms of arguments defined in (A.10), as in

$$S_{\rho>0}(y) = \{(\zeta, \beta) \in \Psi_\rho(y) : \xi(y, \zeta, \beta) \geq \phi(y) + \gamma(\rho)\}, \quad (\text{A.11a})$$

$$S_0(y) = \{(\zeta, \beta) \in \Psi_\rho(y) : \xi(y, \zeta, \beta) > \phi(y)\}. \quad (\text{A.11b})$$

Given that ξ is lower semi-continuous (continuous by A3'), ϕ is upper-semicontinuous (continuous by A3'), Ψ_ρ is dense and pseudo-lower-continuous (lower semicontinuous and ρ -independent), and γ is monotonically increasing; it holds by Proposition 3 of [180] that $\{S_\rho\}_{\rho \geq 0}$ is dense and pseudo-lower-continuous. Corollary (4) then ensures that S_0 is lower semicontinuous. \square

A.2.3 Continuity of Lie Multipliers

We now use the results from Section A.2.2 to prove Theorem 6.4.3.

We begin by recalling the association in (6.24) between the Lie constraint and robust counterpart parameters. Assumptions A1-A4 and A1'-A4' are all active in this section. The parameter set in the Lie setting is $Y = [0, T] \times X \times W$, and the solution set is $Z = \prod_{s=1}^{N_s} K_s^*$ with $\beta = \emptyset$.

A ρ -decreasing family S_ρ (A.9) may be constructed from (6.24). This family has a lower-semicontinuous map S_0 by Theorem A.2.2.

We now review a condition for a continuous selection:

Definition A.2.6. Let $F : Y \rightrightarrows Z$ be a set-valued map. The function $\sigma : Y \rightarrow Z$ is a **selection** for F if $\forall y \in Y : \sigma(y) \in F(y)$.

Theorem A.2.3 (Michael's Theorem, Thm. 9.1.2 of [178]). Let Y be a compact metric space and Z be a Banach Space. If $F : Y \rightrightarrows Z$ has closed convex images for each $y \in Y$, then there exists a continuous selection σ for F .

Note that Michael's Theorem does not require that the images of F in Z should be compact. Michael's theorem requires closed convex images. The images of S_0 are convex, but additional knowledge is needed to ensure that the mappings are closed:

CHAPTER A. APPENDICES

Lemma A.2.4. *Given that the $\mathcal{L}_f v(t, x, w) < 0$ relation holds strictly in the compact space $\Omega = [0, T] \times X \times W$ with $\mathcal{L}_f v$ continuous, the function $\mathcal{L}_f v$ attains its maximal value.*

We now prove Theorem 6.4.3:

Proof. Define τ as $\tau = \max_{(t,x,w) \in \Omega} < 0$ from Lemma A.2.4. All mappings $\{S_\rho\}_{\rho \in [0, \tau]}$ are equal to each other, and therefore both S_τ and S_0 are closed. S_0 satisfies all requirements of Michael's Theorem A.2.3 and therefore has a continuous selection for the Lie multipliers. The below minimal map is one such continuous selection (Prop. 9.3.2 in [178]):

$$m(S_0(t, x, w)) \doteq \left\{ \zeta \in S_0(t, x, w) : \|\zeta\| = \min_{y \in S_0(t, x, w)} \|y\| \right\}. \quad (\text{A.12})$$

□

Remark A.2.3. *We note that the minimum-norm selection A.12 may be used to prove continuity of multipliers when Y is noncompact by Proposition 9.3.2 of [178] and arguments from Theorem 2 of [86]. Problem instances with noncompact Y will introduce conservatism when applying polynomial approximations, given that the Stone-Weierstrass theorem cannot be applied in the noncompact setting.*

A.3 Integral Costs and Robust Counterparts

This appendix discusses Lie nonnegativity constraints with a cost $J(t, x, w)$:

$$\mathcal{L}_f v(t, x, w) + J(t, x, w) \geq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W. \quad (\text{A.13})$$

Constraints with (A.13) appear in OCPs involving integral costs:

$$\inf_w \int_0^T J(t, x(t), w(t)) dt \quad \dot{x}(t) = f(t, x(t), w(t)), \quad x(0) = x_0. \quad (\text{A.14})$$

The expression in (A.13) is an inequality relaxation [7] of the Hamilton-Jacobi-Bellman constraint [107]:

$$\min_{w \in W} \mathcal{L}_f v(t, x, w) + J(t, x, w) = 0 \quad \forall (t, x) \in [0, T] \times X. \quad (\text{A.15})$$

Integral costs with peak estimation are discussed in Equation (5.2) of [6].

This appendix will assume that Assumptions A1-A4 are active.

Representations of (A.13) in standard robust form (6.6) will be worked out for the specific cases of L_∞ , L_1 , and quadratic running costs. Results will be reported as the combination of an extended SDR uncertainty set \tilde{W} (such that $\pi^w \tilde{W} = w$) and terms (a, b) to form (6.6).

A.3.1 L-infinity Running Cost

This subsection will involve a running cost $J(t, x, w) = \|Cw\|_\infty$ for a matrix $C \in \mathbb{R}^{c \times L}$ with $c \geq L$ and $\text{rank}(C) = L$. A new term $\tilde{w} \in \mathbb{R}$ may be introduced to form the lifted SDR uncertainty set

$$\tilde{W}_\infty = \{w \in W, \tilde{w} \in \mathbb{R} : \mathbf{1}_c \tilde{w} - Cw \geq 0, \mathbf{1}_c \tilde{w} + Cw \geq 0\}. \quad (\text{A.16})$$

The weighted L_∞ -running cost Lie term from (A.13) is

$$\begin{aligned} & \mathcal{L}_f v(t, x, w) + \|Cw\|_\infty \geq 0 & \forall (t, x, w) \in [0, T] \times X \times W. \\ & = \mathcal{L}_f v(t, x, w) + \tilde{w} \geq 0 & \forall (t, x, w, \tilde{w}) \in [0, T] \times X \times \tilde{W}_\infty. \end{aligned} \quad (\text{A.17})$$

The correspondence in (6.6) for the L_∞ running cost is (A.17) \tilde{w} by

$$b_0 = \mathcal{L}_{f_0} v(t, x, w) \quad a_0 = 0 \quad (\text{A.18a})$$

$$b_\ell = f_\ell \cdot \nabla_x v(t, x, w) \quad a_\ell = 0 \quad \forall \ell \in 1..L \quad (\text{A.18b})$$

$$b_{L+1} = 1 \quad a_{L+1} = 0. \quad (\text{A.18c})$$

A.3.2 L1 Running Cost

This subsection has a running cost of $J(t, x, w) = \|w\|_1$, as performed by [82]. The standard L_1 lift as reported in [71] introduces $\tilde{w} \in \mathbb{R}^L$ under the constraint

$$\tilde{W}_1 = \{w \in W, \tilde{w} \in \mathbb{R}^L : \tilde{w} - w \geq 0, \tilde{w} + w \geq 0\}. \quad (\text{A.19})$$

The L_1 -running cost Lie term from (A.13) is

$$\begin{aligned} & \mathcal{L}_f v(t, x, w) + \|w\|_1 \geq 0 & \forall (t, x, w) \in [0, T] \times X \times W. \\ & = \mathcal{L}_f v(t, x, w) + \mathbf{1}_L^T \tilde{w} \geq 0 & \forall (t, x, w, \tilde{w}) \in [0, T] \times X \times \tilde{W}_1. \end{aligned} \quad (\text{A.20})$$

The correspondence in (6.6) for the L_1 case (A.20) is:

$$b_0 = \mathcal{L}_{f_0} v(t, x, w) \quad a_0 = 0 \quad (\text{A.21a})$$

$$b_\ell = f_\ell \cdot \nabla_x v(t, x, w) \quad a_\ell = 0 \quad \forall \ell \in 1..L \quad (\text{A.21b})$$

$$b_{\ell'} = 1 \quad a_{\ell'} = 0 \quad \forall \ell' \in L + 1..2L. \quad (\text{A.21c})$$

CHAPTER A. APPENDICES

A.3.3 Quadratic Running Cost

This subsection will discuss the standard convex quadratic cost

$$J(t, x, w) = x^T Px + w^T R w + 2w^T N x \quad P \in \mathbb{S}_+^n, R \in \mathbb{S}_+^L, N \in \mathbb{R}^{L \times n}. \quad (\text{A.22})$$

Let $\Xi = [Q, N^T; N, R]$ be a matrix with factorization $(\Xi^{1/2})^T \Xi^{1/2} = \Xi$. The cone description with mixed quadratic uncertainty is

$$\tilde{W}_2 = \{w \in W, \tilde{w} \in \mathbb{R} : (\Xi^{1/2}[x; w], \tilde{w}, 1/2) \in Q_r^L\}. \quad (\text{A.23})$$

The quadratic-cost Lie expression from (A.13) is

$$\begin{aligned} & \mathcal{L}_f v(t, x, w) + x^T Px + w^T R w + 2w^T N x \geq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W \\ & = \mathcal{L}_f v(t, x, w) + \tilde{w} \geq 0 \quad \forall (t, x, w, \tilde{w}) \in [0, T] \times X \times \tilde{W}_2. \end{aligned} \quad (\text{A.24a})$$

The correspondence in (6.6) for the mixed quadratic case case (A.24a) is

$$b_0 = \mathcal{L}_{f_0} v(t, x, w) \quad a_0 = 0 \quad (\text{A.25a})$$

$$b_\ell = f_\ell \cdot \nabla_x v(t, x, w) \quad a_\ell = 0 \quad \forall \ell \in 1..L \quad (\text{A.25b})$$

$$b_{L+1} = 1 \quad a_{L+1} = 0. \quad (\text{A.25c})$$

Lemma A.3.1. *Lower semicontinuity (Theorem 6.4.3) and convergence (Theorem 6.5.4) is preserved in the mixed set (A.23) (after accounting for the sign changes with the strict $\mathcal{L}_f v > 0$).*

Proof. Let $[\Xi_x^{1/2}, \Xi_w^{1/2}]$ be a column-wise partition of $\Xi^{1/2}$ corresponding to the x and w multiplications. The rotated SOC constraint in (A.23) may be expressed with parameters

$$A_{\text{mix}} = \begin{bmatrix} \Xi_w^{1/2} & 0 \\ \mathbf{0}_{L \times 1} & 1 \\ \mathbf{0}_{L \times 1} & 0 \end{bmatrix} \quad e_{\text{mix}} = \begin{bmatrix} \Xi_x^{1/2} x \\ 0 \\ 1/2 \end{bmatrix} \quad (\text{A.26a})$$

$$G_{\text{mix}} = \emptyset \quad K_{\text{mix}} = Q_r^L \quad (\text{A.26b})$$

forming the conic constraint

$$A_{\text{mix}}[w; \tilde{w}] + e_{\text{mix}} \in K_{\text{mix}}. \quad (\text{A.27})$$

Now consider the assumptions in Section A.2.2. A3' is satisfied because e_{mix} is a continuous (affine) function of x and does not involve t . A4' is also satisfied because A_{mix} is constant in (t, x) and $G_{\text{mix}} = \emptyset$. Assumptions A1-A4 ensure that A1' and A2' are fulfilled, completing the proof. \square

A.4 Polynomial Approximation of the Auxiliary Function

This appendix uses arguments from [6] to prove that Problem (6.16) may be approximated with ε -accuracy by a polynomial auxiliary function. Assumptions A1-A4 are in place, ensuring that $\Omega = [0, T] \times X \times W$ is compact.

Let $\varepsilon > 0$ be an optimality bound, and let $v(t, x) \in C^1([0, T] \times X)$ be an auxiliary function that satisfies constraints (6.16c) and (6.16d) with,

$$\sup_{x \in X_0} v(0, x) \leq d^* + \varepsilon. \quad (\text{A.28})$$

The i -th coordinate of dynamics $\dot{x} = F(t, x, w) = f_0(t, x) + \sum_{\ell=1}^L w_\ell f_\ell(t, x)$ from (6.1) is indexed by $F_i(t, x, w)$.

A tolerance $\eta > 0$ may be chosen as (Equation 4.10 of [6]):

$$\eta < \frac{\varepsilon}{\max(2, 2T, 2T\|F_1\|_{C^0(\Omega)}, \dots, \|F_n\|_{C^0(\Omega)})}. \quad (\text{A.29})$$

A polynomial approximation of the C^1 function v may be performed by Theorem 1.1.2 of [182] to find a polynomial $w \in \mathbb{R}[t, x]$ such that $\|v(t, x) - w(t, x)\|_{C^1(\Omega)} < \eta$ uniformly. The perturbed auxiliary function,

$$V(t, x) = w(t, x) + \varepsilon(1 - t/(2T)), \quad (\text{A.30})$$

satisfies the following strict inequalities from (6.16) (equation 4.12 in [6]),

$$d^* + (5/2)\varepsilon > V(0, x) \quad \forall x \in X_0 \quad (\text{A.31a})$$

$$\mathcal{L}_{F(t,x,w)} V(t, x) < 0 \quad \forall (t, x, w) \in \Omega \quad (\text{A.31b})$$

$$V(t, x) < p(x) \quad \forall (t, x) \in [0, T] \times X. \quad (\text{A.31c})$$

There then exists some finite d such that the polynomial $V(t, x)$ with an optimal solution of at most $d^* + (5/2)\varepsilon$ has degree d [6].

A.5 Polynomial Approximation of Multipliers

Let $(\zeta^c(y), \beta^c(y))$ be a continuous selection of multipliers of $S_0(y)$ (A.9b) (guaranteed to exist by Michael's Theorem A.2.3). This section will prove that there exists a polynomial choice $(\zeta^p(y), \beta^p(y))$ that is also a continuous selection for S_0 .

A.5.1 Continuous Parameterization

Considering the map $G_\rho(y)$ from (A.10d), let $\Phi(y)$ be the matrix inside the $G_\rho(y)$ affine constraint:

$$\Phi = \begin{bmatrix} G^T & \mathbf{0} \\ A^T & a_\bullet^T \end{bmatrix}. \quad (\text{A.32})$$

Assumption A4' imposes that Φ is constant in y . Define H as a constant matrix whose columns span the nullspace of Φ , in which N is the nullity of Φ . Let r be the following least-squares solution (ignoring the conic constraint $\zeta \in K^*$),

$$\theta = \Phi^+[\mathbf{0}; b_\bullet]. \quad (\text{A.33})$$

The vector $\theta(y)$ is a continuous function of y given that b_\bullet is continuous (A3') and Φ is constant.

The set of solutions to $G_\rho(y)$ may be expressed using $(\zeta, \beta) = (\theta + H\psi)$ as

$$G_\rho(y) = \{(\theta + H\psi) \mid \psi \in \mathbb{R}^N, (\theta + H\psi) \in Z\}. \quad (\text{A.34})$$

We will partition (θ, H) according to the resident cones K_s and free values \mathbb{R}^r by

$$\zeta_s = \theta_s + H_s \psi \quad \forall s = 1..N_s \quad (\text{A.35a})$$

$$\beta = \theta_\beta + H_\beta \psi. \quad (\text{A.35b})$$

Given that (ζ^c, β^c) are a continuous selection of (A.9b) satisfying (A.10d), there exists a continuous $\psi_c : Y \rightarrow \mathbb{R}^N$ from (A.34) such that $(\zeta^c, \beta^c) = \theta + H\psi_c$.

A.5.2 Polynomial Approximation

We will use the Stone-Weierstrass theorem to approximate the function $\psi_c : Y \rightarrow \mathbb{R}^n$ by a polynomial vector $\psi_p \in \mathbb{R}[y]^N$ in the compact space Y up to a tolerance $\epsilon > 0$ with

$$\sup_{y \in Y} \|\psi_c(y) - \psi_p(y)\|_\infty \leq \epsilon. \quad (\text{A.36})$$

In order to pose a valid approximation (ζ^p, β^p) , we need to use a notion of centers of cones. We will choose the incenter:

CHAPTER A. APPENDICES

Definition A.5.1 (Def. 2.1 of [183]). Let $(X, \|\cdot\|)$ be a reflexive Banach space with distance $\text{dist}(x_1, x_2) = \|x_1 - x_2\|$, and let S_X be the unit sphere in X . Given a cone $K \subset X$, let $K \cap S_X$ be the set of unit-norm elements of the cone K . The **incenter** of K is the unique solution to

$$\varsigma(K) = \sup_{x \in K \cap S_X} \text{dist}(x; \partial K). \quad (\text{A.37})$$

Remark A.5.1. The following equation lists common cones and their incenters [184]:

$$\mathbb{R}_{\geq 0} : 1 \quad Q^n : (\mathbf{0}_n; 1) \quad \mathbb{S}_+^n : I_n / \sqrt{n}. \quad (\text{A.38a})$$

For a given cone K_s , we define c_s as the incenter of K_s . In the semidefinite case, the incenter will be appropriately vectorized following the vectorial convention of cone containment K_s .

Our approximation (ζ^p, β^p) will be defined using tolerances $\delta_s > 0$ for $s = 1..N_s$:

$$\zeta^p = \theta_s + H_s \psi^p + \delta_s c_s \quad \forall s = 1..N_s \quad (\text{A.39a})$$

$$\beta^p = \theta_\beta + H_\beta \psi^p. \quad (\text{A.39b})$$

The tolerance terms $\delta_s c_s$ will encourage conic containment in K^* . Tolerance terms are therefore unnecessary to encourage containment of the free values $\beta \in \mathbb{R}^r$. Note that (ζ^p, β^p) from (A.39) will be polynomial when the continuous b_\bullet is polynomial (A3).

The approximators (ζ^p, β^p) are related to (ζ^c, β^c) by

$$\zeta_s^p = \zeta_s^c + c_s \delta_s + H_s(\psi^p - \psi^c) \quad (\text{A.40a})$$

$$\beta^p = \beta^c + H_\beta(\psi^p - \psi^c). \quad (\text{A.40b})$$

The term in (A.40a) dominates the worst-case bound

$$\zeta_s^c + c_s \delta_s + H_s(\psi^p - \psi^c) \geq_{K_s^*} \zeta_s^c + c_s^*(\delta_s - \|H_s\|_\infty \epsilon) \quad (\text{A.41})$$

using the Stone-Weierstrass ϵ -approximation (A.36).

A sufficient condition for $\zeta_s^p \in K_s^*$ through (A.41) is that

$$\delta_s \geq \|H_s\|_\infty \epsilon. \quad (\text{A.42})$$

No additional work is needed to ensure $\beta^p \in \mathbb{R}^r$ in (A.40b).

We now move to the strict inequality constraint in (A.9b)

$$e^T \zeta + a_0^T \beta < b_0. \quad (\text{A.43})$$

CHAPTER A. APPENDICES

Substitution of (ζ^p, β^p) from (A.39) into (A.43) leads to

$$\sum_{s=1}^{N_s} e_s^T (\zeta_s^c + c_s \delta_s + H_s(\psi^p - \psi^c)) + a_0^T (\beta^c + H_\beta(\psi^p - \psi^c)) < b_0. \quad (\text{A.44})$$

Recalling that each cone K_s is a subset of the finite-dimensional \mathbb{R}^{n_s} , the left-hand term of (A.44) is upper-bounded using (A.36) by

$$\sum_{s=1}^{N_s} (e_s^T \zeta_s^c + e_s^T c_s \delta_s + \|\text{diag}(e_s) H_s\|_\infty \epsilon n_s) + (a_0^T \beta^c + \|\text{diag}(a_0) H_\beta\|_\infty \epsilon r) \quad (\text{A.45a})$$

$$= \left[\sum_{s=1}^{N_s} e_s^T \zeta_s^c + a_0^T \beta^c \right] + \left[\sum_{s=1}^{N_s} (\delta_s e_s^T c_s + \|\text{diag}(e_s) H_s\|_\infty n_s \epsilon) + \|\text{diag}(a_0) H_\beta\|_\infty \epsilon r \right]. \quad (\text{A.45b})$$

Define Q^* as the finite and positive value

$$Q^* = \min_{y \in Y} b_0 - \left[\sum_{s=1}^{N_s} e_s^T \zeta_s^c + a_0^T \beta^c \right] > 0. \quad (\text{A.46})$$

The minimum in (A.46) is attained because all functions $(b_0, e, a_0, \zeta, \beta)$ are continuous in the compact region Y (just like in Lemma A.2.4).

Successful polynomial-based approximation with $(\zeta^p(y), \beta^p(y)) \in S_0(y)$ will occur if $(\{\delta_s\}, \epsilon)$ are chosen with

$$\epsilon > 0 \quad (\text{A.47a})$$

$$\forall s = 1..N_s : \quad \delta_s \geq \|H_s\|_\infty \epsilon \quad (\text{A.47b})$$

$$\sum_{s=1}^{N_s} (\delta_s e_s^T c_s + \|\text{diag}(e_s) H_s\|_\infty n_s \epsilon) + \|\text{diag}(a_0) H_\beta\|_\infty r \epsilon < Q^*. \quad (\text{A.47c})$$

Shrinking the tolerances $(\{\delta_s\}, \epsilon)$ towards zero will result in approximations of increasing quality. This approximation quality is directly relevant towards establishing convergent bounds in Lie problems, such as in the suboptimal peak estimation task discussed in Appendix A.4.

A.6 Robust Duality and Recovery

This appendix dualizes programs formed by robust Lie constraints (6.25) and forms an interpretation based on occupation measures. It also reviews a technique from [77, 78] to extract approximate polynomial control laws from moment-SOS solutions.

CHAPTER A. APPENDICES

A.6.1 Duality

To simplify explanations, we will consider a polytope-constrained peak estimation problem from (6.36) with $G = 0$ and an set of $W = \{w \in \mathbb{R}^L \mid e - Aw \geq 0\}$. The Lie-robustified peak estimation LP under polytopic uncertainty considered in this appendix is

$$d^* = \inf_{v(t,x), \gamma, \zeta} \gamma \quad (\text{A.48a})$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (\text{A.48b})$$

$$\mathcal{L}_{f_0} v(t, x) + e^T \zeta(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (\text{A.48c})$$

$$-(A^T)_\ell \zeta(t, x) + f_\ell \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \quad (\text{A.48d})$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (\text{A.48e})$$

$$v(t, x) \in C^1([0, T] \times X) \quad (\text{A.48f})$$

$$\zeta_j(t, x) \in C_+([0, T] \times X) \quad \forall j = 1..m. \quad (\text{A.48g})$$

We will derive a weak dual program to (A.48). Define the following measures as multipliers to constraints in (A.48):

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (\text{A.49a})$$

$$\text{Occupation} \quad \mu \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.49b})$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.49c})$$

$$\text{Controlled} \quad \nu \in \mathcal{M}([0, T] \times X) \quad \forall \ell = 1..L \quad (\text{A.49d})$$

$$\text{Constraint-Slack} \quad \hat{\mu}_j \in \mathcal{M}_+([0, T] \times X) \quad \forall j = 1..m. \quad (\text{A.49e})$$

The Lagrangian \mathcal{L} associated with (A.48) is

$$\mathcal{L} = \gamma + \langle v(0, x) - \gamma, \mu_0 \rangle + \langle \mathcal{L}_{f_0} v(t, x) + e^T \zeta(t, x), \mu \rangle + \langle -v(t, x) + p(x), \mu_p \rangle \quad (\text{A.50})$$

$$+ \sum_{j=1}^m \langle -\zeta_j, \hat{\mu}_j \rangle + \sum_{\ell=1}^L \langle f_\ell \cdot \nabla_x v(t, x) - (A^T)_\ell \zeta(t, x), \nu_\ell \rangle$$

$$= \gamma(1 - \langle 1, \mu_0 \rangle) + \langle v(t, x), \delta_0 \otimes \mu_0 \mathcal{L}_{f_0}^\dagger \mu + \sum_{\ell=1}^L (f_\ell \cdot \nabla_x)^\dagger \nu_\ell - \mu_p \rangle \quad (\text{A.51})$$

$$+ \sum_{j=1}^m \langle \zeta_j(t, x), e_j \mu - (\sum_{\ell=1}^L A_{j\ell} \nu_\ell) - \hat{\mu}_j \rangle + \langle p(x), \mu_p \rangle. \quad (\text{A.52})$$

CHAPTER A. APPENDICES

The dual measure LP of (A.48) is

$$p^* = \sup_{(A.49)} \inf_{\gamma, v, \zeta} \mathcal{L} \quad (\text{A.53a})$$

$$= \sup \langle p(x), \mu_p \rangle \quad (\text{A.53b})$$

$$\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_{f_0}^\dagger \mu + \sum_{\ell=1}^L (f_\ell \cdot \nabla_x)^\dagger \nu_\ell \quad (\text{A.53c})$$

$$e_j \mu = (\sum_{\ell=1}^L A_{j\ell} \nu_\ell) + \hat{\mu}_j \quad \forall j = 1..m \quad (\text{A.53d})$$

$$\langle 1, \mu_0 \rangle = 1 \quad (\text{A.53e})$$

$$\text{Measures from (A.49).} \quad (\text{A.53f})$$

Remark A.6.1. Program (A.53) should be compared against the standard peak estimation program (3.2). Constraint (A.53c) is a robustified Liouville equation. Constraint (A.53d) is a sequence of domination conditions, as detailed in Section 2.2.3. Constraint (A.53e) enforces that μ_0 is a probability measure.

Lemma A.6.1. The measure program A.53 upper-bounds on (6.14) with $p^* \geq P^*$.

Proof. Let $t^* \in (0, T]$ be a stopping time of a trajectory of (6.1) with applied control input $w(t)$ starting from an initial condition $x_0 \in X_0$. Measures from (A.53) may be constructed from the trajectory $x(t | x_0, w(\cdot))$.

The probability measures are $\mu_0 = \delta_{x=x_0}$ and $\mu_p = \delta_{t=t^*, x=x(t^*|x_0, w(\cdot))}$. Relaxed occupation measures may be chosen as the occupation measures of the following evaluation maps in the times $t \in [0, t^*]$:

$$\mu : t \mapsto (t, x(t | x_0, w(\cdot))) \quad (\text{A.54a})$$

$$\nu_\ell : t \mapsto (t, w_\ell(t)x(t | x_0, w(\cdot))) \quad \forall \ell = 1..L \quad (\text{A.54b})$$

$$\hat{\mu}_j : t \mapsto (t, (e_j - A_j w(t))x(t | x_0, w(\cdot))) \quad \forall j = 1..m. \quad (\text{A.54c})$$

Every trajectory $x(t | x_0, w(\cdot))$ has a feasible measure representation, proving the upper-bounding theorem. \square

Theorem A.6.2. Strong duality holds between (A.53) with (A.48) $d^* = p^* = P^*$ under assumptions A1-A6.

Proof. The bound $d^* \geq p^*$ holds by weak duality [75, 185]. Lemma A.6.1 proves that $p^* \geq P^*$, together forming the chain $d^* \geq p^* \geq P^*$. Theorem 6.4.2 proves that the optimal value d^* from the

CHAPTER A. APPENDICES

robust (A.48) equals the optimal value of the non-robust (6.16), which is in turn equal to P^* from (6.14) by Theorem 2.1 of [7]. Since d^* and P^* are equal, it holds that the sandwiched p^* satisfies $d^* = p^* = P^*$. \square

Lemma A.6.3. *Under A1-A6 and the further assumption that the polytope W is compact all of the nonnegative measures in (A.53) are bounded.*

Proof. Boundedness of a nonnegative measure will be demonstrated by showing that the measure has finite mass and it is supported on a compact set. Assumptions A1-A2 posit compactness of $[0, T] \times X$. The probability measures are $\langle 1, \mu_0 \rangle = 1$ (by (A.53e)), $\langle 1, \mu_p \rangle = 1$ (by (A.53c) with $v(t, x) = 1$). The relaxed occupation measure μ is bounded with $\langle 1, \mu \rangle \leq T$ (by (A.53c) under A1).

Applying a test function $\zeta_j = 1$ to the domination constraint (A.53d) leads to

$$e\langle 1, \mu \rangle = A\langle 1, \nu \rangle + \langle 1, \hat{\mu} \rangle \Rightarrow e\langle 1, \mu \rangle \geq A\langle 1, \nu \rangle. \quad (\text{A.55})$$

where the measure pairings are vectorized for convenience. The pairings $\langle 1, \nu \rangle$ are members of the $\langle 1, \mu \rangle$ -scaled compact polytope W , proving that the constraint-slack measures $\hat{\mu}_j$ is bounded for each $j \in 1..m$. \square

Remark A.6.2. *The signed measures ν_ℓ in (A.53) have unbounded TV norm. Each signed measure $\nu_\ell \in \mathcal{M}([0, T] \times X)$ can be decomposed into nonnegative measures by a Hahn-Jordan decomposition,*

$$\nu_\ell = \nu_\ell^+ - \nu_\ell^-, \quad \nu_\ell^+, \nu_\ell^- \in \mathcal{M}_+([0, T] \times X) \quad \forall \ell = 1..L \quad (\text{A.56})$$

$$\nu_\ell^+ \perp \nu_\ell^- \quad \forall \ell = 1..L. \quad (\text{A.57})$$

A direct substitution of (A.56) into (A.53) will leave the TV norm $\|\nu_\ell\|_{TV} = \langle 1, \nu_\ell^+ + \nu_\ell^- \rangle$ as a possibly unbounded degree of freedom, because only the mass of the difference $\langle 1, \nu_\ell^+ - \nu_\ell^- \rangle$ is constrained in (A.53d). Under the assumption that the SDR set W is compact, the measures ν_ℓ^+ and ν_ℓ^- may be bounded by adding new mass constraints to (A.53)

$$M_\ell^+ = \max_{w \in W, w_\ell \geq 0} w_\ell, \quad \langle 1, \nu_\ell^+ \rangle \leq \langle 1, \mu \rangle M_\ell^+ \quad \forall \ell = 1..L \quad (\text{A.58a})$$

$$M_\ell^- = \min_{w \in W, w_\ell \leq 0} w_\ell, \quad \langle 1, \nu_\ell^- \rangle \leq \langle 1, \mu \rangle M_\ell^- \quad \forall \ell = 1..L. \quad (\text{A.58b})$$

The addition of (A.58) will not change the optimum value p^* of (A.53). However, the dual of (A.53) with constraints in (A.58) will be different from (A.48), and will no longer feature equality constraints in (A.48d).

CHAPTER A. APPENDICES

Remark A.6.3. *The duality results of this appendix subsection may be extended to other compact SDR sets W . Special caution must be taken notationally when referring to the adjoints of affine maps (from (6.5)) and the dual spaces of cone-valued continuous functions $C([0, T] \times X \rightarrow K_s^*)'$.*

Remark A.6.4. *The work in [77] performs optimal control in the unit box $W = [-1, 1]^L$, resulting in measure programs that have the form of (A.56) (excluding the orthogonality constraint (A.57)) with an additional bounded-mass constraint:*

$$\nu_\ell^+ + \nu_\ell^- + \hat{\mu}_\ell = \mu_\ell \quad \nu_\ell^+ - \nu_\ell^- = \nu_\ell \quad \forall \ell = 1..L. \quad (\text{A.59})$$

The work in [78] rescales the dynamics to ensure that $W = [0, 1]^L$. The measure ν_ℓ^- can be set to zero in the nonnegative box case, and the problem involves only nonnegative measures with $\nu_\ell = \nu_\ell^+$ for each ℓ .

A.6.2 Recovery

Let (v, ζ) be a degree- $2k$ solution to the SOS program (6.27) associated with (A.48). Let Q_0 be the solved Gram matrix associated with the Lie constraint (A.48c) (SOS constraint (6.27a)), and let σ_ℓ be the vector of dual variables corresponding to the equality constraint (A.48d) (finite-degree (6.27b)). By strong duality in the hierarchy ([27] and extensions from [97, Theorem 4 and Lemma 4]), the SDP dual variable to Q_0 is the moment matrix $\mathbb{M}_d[\mathbf{m}]$ in which \mathbf{m} is a moment sequence of μ . Similarly, the dual variables of σ_ℓ are moment sequences \mathbf{m}_ℓ of a signed measure μ_ℓ for each $\ell = 1..L$ (because every symmetric matrix may be expressed as the difference between two PSD matrices).

An approximate control law $\hat{w}_\ell(t, x)$ for all $\ell = 1..L$ may be recovered from the degree- $\leq 2k$ moments of \mathbf{m} and the degree- $\leq k$ moments of \mathbf{m}_ℓ (written as $\mathbf{m}_\ell^{\leq d}$) by [77, Equation 41]

$$\mathbf{y}_\ell = \mathbb{M}_d \mathbf{m}^{-1}(\mathbf{m}_\ell^{\leq d}) \quad \hat{w}_\ell(t, x) = \sum_{(\alpha, \beta) \in \mathbb{N}^{L+1}, |\alpha| + \beta \leq d} \mathbf{y}_{\ell \alpha \beta} x^\alpha t^\beta. \quad (\text{A.60a})$$

Controllers from (A.60) will converge in an L_1 sense to the optimal control law by [77, Theorem 8] as the degree increases. It remains an open problem to quantify performance indices when deploying finite-degree recovered controls on system (6.1).

A.7 Strong Duality of Linked Semidefinite-Measure Programs

The work in [23] gives sufficient condition to ensure strong duality in the framework of linear programming on measures. This appendix generalizes the results of [23] by forming a framework of convex programming on measures with infinite-dimensional linear constraints and finite dimensional LMI constraints on moments. In particular, we add the case of optimization over Borel measures with SOC constraints on their moments to the original framework [23].

Let $M, m_0 \in \mathbb{N}$ be positive integers. For $i = 1..M$, let $m_i \in \mathbb{N}$ and $X_i \subset \mathbb{R}^{m_i}$ be a compact set. Let:

- $\mathcal{X}_0 \subset \mathbb{S}^{m_0}$ be a vector space of symmetric matrices. Specific instances of \mathcal{X}_0 could be $\{\text{diag}(\chi) \mid \chi \in \mathbb{R}^{m_0}\}$ (the space of diagonal matrices, corresponding to linear programming), or $\mathcal{X}_0 = \mathbb{S}^{m_0}$ (the space of all symmetric matrices, corresponding to semidefinite programming). In particular, there exists such a space to represent second order cone programming [89]),
- $\mathcal{X}_\infty = \mathcal{M}(X_1) \times \dots \times \mathcal{M}(X_M)$ be a vector space of signed Borel measures, equipped with its weak-* topology, so that its topological dual is $\mathcal{X}_\infty^* = C(X_1) \times \dots \times C(X_M)$,
- $\mathcal{X} = \mathcal{X}_0 \times \mathcal{X}_\infty$ be our decision space, with topological dual $\mathcal{X}^* = \mathcal{X}_0^* \times \mathcal{X}_\infty^*$,
- A Banach space \mathcal{Y} with dual \mathcal{Y}^* that will represent our constraint space for equality constraints. In the context of the moment-SOS hierarchy, \mathcal{Y} is chosen as a product space of smooth/polynomial functions defined on compact sets,
- $\mathcal{X}_+ = \{(X, \mu_1, \dots, \mu_M) \in \mathcal{X} \mid X \succeq 0, \quad \forall i = 1..M, \mu_i \in \mathcal{M}_+(X_i)\}$ and $\mathcal{X}_+^* = \{(Y, v_1, \dots, v_M) \in \mathcal{X}^* \mid Y \succeq 0, \quad \forall i = 1..M, v_i \geq 0\}$ be two convex cones.

For $\phi = (X, \mu_1, \dots, \mu_M) \in \mathcal{X}$ and $\psi = (Y, v_1, \dots, v_M) \in \mathcal{X}^*$, we define the duality

$$\langle \psi, \phi \rangle_{\mathcal{X}} = \text{Tr}(X Y) + \sum_{i=1}^M \int_{X_i} v_i(x_i) d\mu_i(x_i). \quad (\text{A.61})$$

Similarly, we denote by $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$ the duality between elements of \mathcal{Y} and \mathcal{Y}^* . Let $A : \mathcal{X} \rightarrow \mathcal{Y}^*$ be a continuous linear map, $y \in \mathcal{Y}^*$ be a vector of continuous linear forms (when \mathcal{Y} is made of polynomials, y is a moment sequence), $C \in \mathcal{X}_0$, $p \in \mathbb{R}[x_1] \times \dots \times \mathbb{R}[x_M] \subset \mathcal{X}_\infty^*$ be a vector of

CHAPTER A. APPENDICES

polynomials $\gamma = (C, g) \in \mathcal{X}^*$. We consider the following moment-SDP problem:

$$\begin{aligned} p_M^* &= \sup_{\phi \in \mathcal{X}_+} \langle \gamma, \phi \rangle_{\mathcal{X}} \\ A\phi &= y \end{aligned} \tag{A.62a}$$

with dual problem

$$\begin{aligned} d_M^* &= \inf_{w \in \mathcal{Y}} \langle w, y \rangle_{\mathcal{Y}} \\ A^\dagger w - \gamma &\in \mathcal{X}_+^*. \end{aligned} \tag{A.62b}$$

It is well known that weak duality $p_M^* \leq d_M^*$ always holds [185]. In this section, we will prove that under some mild assumptions, strong duality $p_M^* = d_M^*$ also holds.

We first prove a simple lemma on strong duality conditions.

Lemma A.7.1. *In this lemma, we consider again the duality pair (A.62), but with generic spaces \mathcal{X} , \mathcal{Y} and a convex cone \mathcal{X}_+ . We also define, for any vector space \mathcal{Z} containing some vector η and any linear map $U : \mathcal{X} \rightarrow \mathcal{Z}$, the level set $U_\eta = \{\phi \in \mathcal{X} \mid U\phi = \eta\}$. In such setting, we assume that*

A1' $\exists \phi \in \mathcal{X}_+$ such that $A\phi = y$.

A2' $A_0 \cap \gamma_0 \cap \mathcal{X}_+ = \{0\}$.

A3' $\exists \psi \in \mathcal{X}^$ such that*

- (a) $\langle \psi, \mathcal{X}_+ \rangle_{\mathcal{X}} \subset \mathbb{R}_+$
- (b) $\psi_0 \cap \mathcal{X}_+ = \{0\}$
- (c) $\psi_1 \cap \mathcal{X}_+$ is compact.

Then, $p_M^ = d_M^*$.*

Moreover, if $p_M^ < \infty$, then there is an optimal ϕ^* feasible for (A.62a) such that $\langle \gamma, \phi^* \rangle_{\mathcal{X}} = p_M^*$.*

Proof. We use [185, Chap. IV: Thm (7.2), Lem (7.3)]. Consider the cone

$$\mathcal{X}_A^\gamma = \{(A\phi, \langle \gamma, \phi \rangle_{\mathcal{X}}) \mid \phi \in \mathcal{X}_+\}.$$

CHAPTER A. APPENDICES

Theorem (7.2) of [185] ensures that under A1' and closedness of \mathcal{X}_A^γ , strong duality holds, and that $p_M^* < \infty$ then implies existence of an optimal ϕ^* . Then, Lemma (7.3) of [185] states that if \mathcal{X}_+ has a compact convex base and A2' holds, then $A_{\mathcal{X}}^\gamma$ is closed. Thus, we need to find a compact convex base of \mathcal{X}_+ .

Let $\phi \in \mathcal{X}_+ \setminus \{0\}$. A3'.(a)-(b) ensure that $\langle \psi, \phi \rangle_{\mathcal{X}} > 0$ so that $\tilde{\phi} = \frac{\phi}{\langle \psi, \phi \rangle_{\mathcal{X}}}$ is well defined and belongs to the cone $\mathcal{X}_+ \setminus \{0\}$. Moreover, $\langle \psi, \tilde{\phi} \rangle_{\mathcal{X}} = 1$ is clear by definition, so that $\tilde{\phi} \in \psi_1 \cap \mathcal{X}_+$. This proves that any $\phi \in \mathcal{X}_+ \setminus \{0\}$ can be described as $\phi = \langle \psi, \phi \rangle_{\mathcal{X}} \tilde{\phi}$ with $\tilde{\phi} \in \psi_1 \cap \mathcal{X}_+$ and $\langle \psi, \phi \rangle_{\mathcal{X}} > 0$, which is the definition of $\psi_1 \cap \mathcal{X}_+$ being a base of \mathcal{X} . By compactness assumption A3'.(c), we deduce that the assumptions of Lemma (7.3) of Theorem (7.2) of [185] hold: X_A^γ is closed and thus $p_M^* = d_M^*$. \square

Theorem A.7.2. *Suppose that there exists $B > 0$ such that for all $\phi = (X, \mu_1, \dots, \mu_M)$ feasible for (A.62a), one has $\text{Tr}(X^2) \leq B^2$ and for $i = 1..M$, $\langle 1, \mu_i \rangle \leq B$. Also assume that at least one such feasible ϕ exists. Then, $p_M^* = d_M^*$. Moreover, there exists an optimal ϕ^* such that $A\phi^* = y$ and $\langle \gamma, \phi^* \rangle_{\mathcal{X}} = p_M^*$.*

Proof. We prove that the assumptions of Lemma A.7.1 hold. First of all, \mathcal{X}_+ is indeed a convex cone, as it is the product of convex cones $\mathbb{S}_+^{m_0}$ and $\mathcal{M}_+(X_i)$ under A1'.

Next, we focus on hypothesis A2' Let $\phi = (X, \mu_1, \dots, \mu_M) \in A_0 \cap \gamma_0 \cap \mathcal{X}_+$. We want to prove that $\phi = 0$. Let $\phi^{(0)} = (X^{(0)}, \mu_1^{(0)}, \dots, \mu_M^{(0)}) \in \mathcal{X}_+$ such that $A\phi^{(0)} = y$. Define, for $t \geq 0$, $\phi^{(t)} = \phi^{(0)} + t\phi$. Let $t \geq 0$. Since \mathcal{X}_+ is a convex cone, $\phi^{(t)} \in \mathcal{X}_+$. In addition,

$$A\phi^{(t)} = A\phi^{(0)} + tA\phi = A\phi^{(0)} = y,$$

so that $\phi^{(t)}$ is feasible for (A.62a). Thus, by assumption,

$$\begin{aligned} B^2 &\geq \text{Tr}((X^{(0)} + tX)^2) \\ &= \text{Tr}(X^{(0)2} + 2tX^{(0)}X + t^2X^2) \\ &= \text{Tr}(X^{(0)2}) + 2t\text{Tr}(X^{(0)}X) + t^2\text{Tr}(X^2) \\ &= t^2\text{Tr}(X^2) + \underset{t \rightarrow \infty}{o}(t^2). \end{aligned}$$

Staying bounded when t goes to infinity requires $\text{Tr}(X^2) = 0$, implying that $X = 0$. The same reasoning replacing $\text{Tr}(X^2)$ with $\langle 1, \mu_i \rangle$ yields that for all $i = 1..M$, $\mu_i = 0$. Thus, $\phi = 0$ and A2' holds.

CHAPTER A. APPENDICES

We turn to A3' and consider $\psi = (I_{m_0}, \mathbf{1}_M)$. Note that if $(X, \mu_1, \dots, \mu_M) \in \mathcal{X}_+$, then $\text{Tr}(X) \geq 0$ and $\forall i = 1..M : \langle 1, \mu_i \rangle \geq 0$ (i.e. $\langle \psi, \mathcal{X}_+ \rangle_{\mathcal{X}} \subset \mathbb{R}_+$). Moreover, the equality cases in those inequalities only hold for $X = 0$ and $\mu_i = 0$ respectively, so that $\psi_0 \cap \mathcal{X}_+ = \{0\}$. It only remains to prove that $\psi_1 \cap \mathcal{X}_+$ is compact.

First, it is bounded for the norm $\|(X, \mu_1, \dots, \mu_M)\| = \sqrt{\text{Tr}(X^2)} + \sum_{i=1}^M \|\mu_i\|_{TV}$ where the TV norm (2.1) is

$$\|\mu\|_{TV} = \sup\{\langle v, \mu \rangle \mid -1 \leq v \leq 1\}.$$

In particular, the TV norm of a nonnegative measure $\mu \in \mathcal{M}_+(X)$ is equal to its mass $\|\mu\|_{TV} = \langle 1, \mu \rangle$.

Indeed, let $(X, \mu_1, \dots, \mu_M) \in \psi_1 \cap \mathcal{X}_+$; then $\text{Tr}(X) \leq 1$ and for $i = 1..M, 1 \geq \langle 1, \mu_i \rangle = \|\mu_i\|_{TV}$. As X is positive semidefinite, $\text{Tr}(X) \leq 1$ means that none of the eigenvalues of X are bigger than 1, so that $\text{Tr}(X^2) \leq m_0$, as it is the sum of the squares of the eigenvalues of X . Thus, for all $\phi \in \psi_1 \cap \mathcal{X}_+$, $\|\phi\| \leq m_0 + M$.

Then, it is also closed for the weak-* topology of \mathcal{X} by continuity of $\langle \psi, \cdot \rangle_{\mathcal{X}}$ and closedness of \mathcal{X}_+ as the product of closed sets $\mathbb{S}_+^{m_0}$ and $\mathcal{M}_+(X_i)$. Thus, $\psi_1 \cap \mathcal{X}_+$ is weak-* closed and bounded: according to the Banach-Alaoglu theorem, it is compact: assumption A3' of Lemma A.7.1 holds, which concludes the proof of strong duality.

Finally, let $\phi = (X, \mu_1, \dots, \mu_M) \in \mathcal{X}_+$ feasible for (A.62a). Then, by assumption, $\text{Tr}(X^2) \leq B^2$ and $\langle 1, \mu_i \rangle \leq B$ for all $i = 1..M$. Thus, one has

$$\begin{aligned} \langle \gamma, \phi \rangle_{\mathcal{X}} &= \text{Tr}(C X) + \sum_{i=1}^M \langle g_i, \mu_i \rangle \\ &\leq \sqrt{\text{Tr}(C^2) \text{Tr}(X^2)} + \sum_{i=1}^M \langle g_i, \mu_i \rangle \quad \text{using the Cauchy-Schwarz inequality} \\ &\leq \sqrt{\text{Tr}(C^2) \text{Tr}(X^2)} + \sum_{i=1}^M \sup_{X_i} |g_i| \langle 1, \mu_i \rangle \\ &\leq \left(\sqrt{\text{Tr}(C^2)} + \sum_{i=1}^M \|g_i\|_{\infty} \right) B \end{aligned}$$

so that taking the supremum over all feasible ϕ yields

$$p_M^* \leq \left(\sqrt{\text{Tr}(C^2)} + \sum_{i=1}^M \|g_i\|_{\infty} \right) B < \infty,$$

which is the last hypothesis of Lemma A.7.1 and ensures existence of an optimal solution. \square

A.8 Strong Duality of Chance-Peak Linear Programs

In order to apply the strong duality results of Appendix A.7 towards the chance-peak problem (Theorem 10.3.7), we need to provide an SDP-representation of the SOC cone.

Lemma A.8.1. *An element $(\kappa, \lambda) \in Q^n$ satisfies the LMI [89]*

$$\Phi = \begin{bmatrix} \lambda & \kappa^T \\ \kappa & \lambda I_n \end{bmatrix} \succeq 0. \quad (\text{A.63})$$

Proof. When $\lambda = 0$, the containment $(\kappa, 0) \in Q^n$ requires that $\kappa = \mathbf{0}_n$. The matrix in Φ (A.63) is therefore the $\mathbf{0}_{n \times n}$ matrix which is PSD. Now consider the case where $\lambda > 0$. A Schur complement of Φ yields the constraint $\lambda - \kappa^T \kappa / \lambda \geq 0$. Multiplying through by the positive λ results in $\lambda^2 - \|\kappa\|_2^2 \geq 0$, $\lambda > 0$, which is the definition of the SOC cone $(\kappa, \lambda) \in Q^n$. Lemma A.8.1 is therefore proven. \square

Lemma A.8.1 ensures that problem (10.19) is an instance of the more generic (A.62a) with $M = 2$, $X_1 = X_2 = [0, T] \times X$, $\mathcal{Y} = C^2([0, T] \times X) \times \mathbb{R}^3$ and

$$\mathcal{X}_0 = \left\{ \begin{bmatrix} \lambda & \kappa^T \\ \kappa & \lambda I_3 \end{bmatrix} \mid \lambda \in \mathbb{R}, \kappa \in \mathbb{R}^3 \right\}.$$

Therefore, we only need to verify that the hypotheses of Theorem A.7.2 hold in our specific case.

Letting $q = ([q_1, q_2, q_3], q_4) \in Q^3$ be an SOC-constrained variable, we define the matrix Φ from (A.63) as

$$\Phi = \begin{bmatrix} q_4 & q_1 & q_2 & q_3 \\ q_1 & q_4 & 0 & 0 \\ q_2 & 0 & q_4 & 0 \\ q_3 & 0 & 0 & q_4 \end{bmatrix}. \quad (\text{A.64})$$

Theorem A.7.2 requires the following sufficient conditions to prove strong duality between (10.19) and (10.20) and their optimality obtainment.

R1 There exists a feasible solution for (μ_τ, μ, q) from (10.19).

R2 The measures μ_τ, μ are bounded.

R3 The square of the matrix Φ from (A.64) has bounded trace.

CHAPTER A. APPENDICES

We start with R1. Letting $x(t \mid x_0)$ be an SDE trajectory from (10.4) and $t^* \in (0, T]$ be a stopping time, we define μ as the occupation measure of $x(t \mid x_0)$ and μ_{t^*} as its time- t^* state distribution μ_{t^*} . Feasible choices for entries of the SOC-constrained q are (from Lemma 10.3.4)

$$q_1 = 1 - \langle p^2, \mu_{t^*} \rangle, \quad q_2 = \sqrt{\langle p^2, \mu_{t^*} \rangle - \langle p, \mu_{t^*} \rangle^2}, \quad q_3 = 2\langle p, \mu_{t^*} \rangle, \quad q_4 = 1 + \langle p^2, \mu_{t^*} \rangle. \quad (\text{A.65})$$

Requirement R2's satisfaction follows the statement in Lemma 10.4.2 that μ_{t^*}, μ are bounded under A1-A3.

We end with R3. The trace $\text{Tr}(\Phi^2) = \sum_{ij} \Phi_{ij}^2$ is equal to

$$\text{Tr}(\Phi^2) = 2q_1^2 + 2q_2^2 + 2q_3^2 + 4q_4^2 \quad (\text{A.66a})$$

$$= 2(1 - \langle p^2, \mu_{t^*} \rangle)^2 + 2(\sqrt{\langle p^2, \mu_{t^*} \rangle - \langle p, \mu_{t^*} \rangle^2})^2 + 2(2\langle p, \mu_{t^*} \rangle)^2 + 4(1 + \langle p^2, \mu_{t^*} \rangle)^2 \quad (\text{A.66b})$$

$$= 6 + 6\langle p, \mu_{t^*} \rangle^2 + 6\langle p^2, \mu_{t^*} \rangle + 6\langle p^2, \mu_{t^*} \rangle^2. \quad (\text{A.66c})$$

Let $\Pi_1 = \max_{x \in X} p(x)$ and $\Pi_2 = \max_{x \in X} p(x)^2$ be bounds on p and p^2 in X . Both Π_1 and Π_2 will be finite by the compactness of X (A1) and the continuity of p within X (A3). Given that μ_{t^*} is a probability distribution supported in X , the moments of μ_{t^*} will be bounded by $\langle p, \mu_{t^*} \rangle \leq \Pi_1$ and $\langle p^2, \mu_{t^*} \rangle \leq \Pi_2$. The squared-trace in (A.66) can be upper-bounded by a finite value B such that

$$\text{Tr}(\Phi^2) \leq 6(1 + \Pi_1^2 + \Pi_2 + \Pi_2^2) = B < \infty. \quad (\text{A.67})$$

The finite bound $B \in [0, \infty)$ from (A.67) validates R3, and completes all conditions necessary for Theorem A.7.2 to provide for strong duality and optima attainment.

A.9 Delay Structures

This chapter has focused on supremizing $p(x)$ in (11.1) over continuous-time systems with a discrete delay $x(t - \tau)$. This subsection will discuss peak estimation of $p(x)$ with respect to other types of dynamics and delay structures.

A.9.1 Proportional Time-Delays

A system with a proportional delay is defined with respect to a scaling term $\kappa \in [0, 1)$ as

$$\dot{x}(t) = f(t, x(t), x(\kappa t)). \quad (\text{A.68})$$

CHAPTER A. APPENDICES

Proportional time delays are observed in the current collection of a pantograph on a streetcar [186]. References on functional differential equation with proportional time delay include [187, 188, 189, 190]. MATLAB uses the command `ddesd` to solve DDEs with time-dependent delays by an RK4 algorithm [191]. Other numerical algorithms specifically for proportional delays include [192, 193, 194].

The peak estimation problem over (A.68) is

$$P^* = \sup_{t^* \in [0, T], x_0 \in X_0} p(x(t^* | x_0)) \quad (\text{A.69a})$$

$$\dot{x} = f(t, x(t), x(\kappa t)) \quad \forall t \in [0, T]. \quad (\text{A.69b})$$

An MV-solution for proportional time delays is

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (\text{A.70a})$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.70b})$$

$$\text{Time-Slack} \quad \nu \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.70c})$$

$$\text{Occupation Start} \quad \bar{\mu}_0 \in \mathcal{M}_+([0, \kappa T] \times X^2) \quad (\text{A.70d})$$

$$\text{Occupation End} \quad \bar{\mu}_1 \in \mathcal{M}_+([\kappa T, T] \times X^2). \quad (\text{A.70e})$$

Note how the MV-solution (A.70) lacks a history measure μ_h as compared with (11.3), and also how the limits on (A.70d)-(A.70e) are $[0, \kappa T]$ and $[\kappa T, T]$ respectively.

The Lie derivative operator \mathcal{L} with $\mathcal{L}v = (\partial_t + f(t, x_0, x_1) \cdot \nabla_{x_0})v(t, x_0)$ is the same as in the discrete-delay case (11.5) but under dynamics (A.68).

The consistency constraint follows from a modification of Lemma 11.3.1:

Lemma A.9.1. *Let $x(\cdot)$ be a solution to (A.69b) with an initial condition of $x_0 \in X_0$ and a stopping time of $t^* \in [0, T]$. The following pairs of integral are equal for all $\phi \in C([0, T] \times X)$:*

$$\int_0^{t^*} \phi(t, x(\kappa t)) dt = \frac{1}{\kappa} \int_0^{\min(t^*/\kappa, T)} \phi(t'/\kappa, x(t)) dt'. \quad (\text{A.71})$$

Proof. This relation is due to a change of variable with $t' \leftarrow \kappa t$. □

The resultant consistency constraint w.r.t. the measures in (A.70) is

$$\langle \phi(t, x_1), \bar{\mu}_0 + \bar{\mu}_1 \rangle + \langle \phi(t, x), \nu \rangle = \langle \phi(t/\kappa, x_0)/\kappa, \bar{\mu}_0 \rangle. \quad (\text{A.72})$$

CHAPTER A. APPENDICES

Expressing the linear expansion operator E_κ as $E_\kappa \phi(t, x) = \phi(t/\kappa, x_0)\kappa$, the measure LP for problem (A.69) is

$$p^* = \sup \langle p, \mu_p \rangle \quad (\text{A.73a})$$

$$\langle 1, \mu_0 \rangle = 1 \quad (\text{A.73b})$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_{\#}^{tx_0} \mathcal{L}_f^\dagger (\bar{\mu}_0 + \bar{\mu}_1) \quad (\text{A.73c})$$

$$\pi_{\#}^{tx_1} (\bar{\mu}_0 + \bar{\mu}_1) + \nu = E_{\#}^\kappa (\pi_{\#}^{tx_0} \bar{\mu}_0) \quad (\text{A.73d})$$

$$\text{Measure Definitions from (A.70).} \quad (\text{A.73e})$$

Problem (A.73) upper-bounds (A.69) by following the reasoning from Theorem 11.3.2 for the proportional-delay case.

Remark A.9.1. *Proportional and discrete delays can be applied together to form dynamics*

$$\dot{x}(t) = f(t, x(t), x(\kappa t - \tau)). \quad (\text{A.74})$$

Causality of (A.74) requires that $\kappa \in [0, 1)$ and $\tau \geq 0$. A consistency constraint may be posed using an integral relation

$$\int_0^T \phi(t, x(\kappa t - \tau)) dt = \int_{-\tau}^{\kappa T - \tau} \phi((t + \tau)/\kappa, x(t))/\kappa dt, \quad (\text{A.75})$$

used as a step towards forming Lemmas 11.3.1 and A.9.1.

A.9.2 Discrete-Time Systems

This subsection will concentrate on a discrete-time system with a long time delay. The discrete-time system $x[t]$ is defined w.r.t. a delay $\tau \in \mathbb{N}$, and a time horizon $T \in \mathbb{N}$ under the assumption that $\tau < T$.

The peak estimation program for a system with discrete-time dynamics and one time delay τ is

$$P^* = \sup_{t^* \in 0..T, x_h[\cdot]} p(x[t | x_h]) \quad (\text{A.76a})$$

$$\dot{x} = f(t, x[t], x[t - \tau]) \quad \forall t \in 1..T \quad (\text{A.76b})$$

$$x[t] = x_h[t] \quad \forall t \in [-\tau, 0] \quad (\text{A.76c})$$

$$x_h[\cdot] \in \mathcal{H}. \quad (\text{A.76d})$$

CHAPTER A. APPENDICES

Delayed dynamics (A.76b) may be implemented as a non-delayed discrete system by state inflation in terms of $x[t - (0..\tau)]$ [158]. Such state augmentation could lead to a large number of variables in systems analysis and result in intractably large computational problems.

This subsection will define an MV-solution using the variables from (11.3), in which the measures with the maximum number of variables are $(\bar{\mu}_0(t, x_0, x_1), \bar{\mu}_1(t, x_0, x_1))$.

A.9.2.1 History-Validity

The history-validity constraint for discrete-time systems will separate the history $x_h[t]$ into a time-zero component (μ_0) and a history component $t \in -\tau..-1$ (μ_h). The time-zero component is $\mu_0 \in X_0$, as in the 11.3.2.1.

The history measure μ_h should represent a history $x_h[t]$ defined between $t \in -\tau..-1$. This may be imposed by setting the t -marginal of μ_h to a train of Dirac-deltas supported at sample times $-\tau..-1$ as

$$\pi_{\#}^t \mu_h = \sum_{t=-\tau}^{-1} \delta_t. \quad (\text{A.77})$$

A.9.2.2 Liouville

The discrete-time Liouville equation 4.3a applied to the dynamics (A.76b) for all test functions $v \in C([0, T + 1] \times X)$ is

$$\langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle v(t + 1, f(t, x_0, x_1) - v(t, x_0, x_1), \bar{\mu}_0 + \bar{\mu}_1 \rangle. \quad (\text{A.78})$$

The Liouville constraint in (A.78) will be abbreviated (using the identity operator $Id(x) = x$) as

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_{\#}^{tx_0}((t + 1, f)_{\#} - Id_{\#})(\bar{\mu}_0 + \bar{\mu}_1). \quad (\text{A.79})$$

A.9.2.3 Consistency

The consistency constraint for dynamics (A.76b) may be derived from the following Lemma,

Lemma A.9.2. *Let $x[\cdot]$ be a trajectory of (A.76b) given an initial history x_h and a stopping time of $t^* \in 0..T$. It follows that the below pair of summations are equal for all $\phi \in C([0, T] \times X)$:*

$$\left(\sum_{t=0}^{t^*} + \sum_{t=t^*}^{\min(T, t^* + \tau)} \right) \phi(t, x[t - \tau]) dt = \sum_{t'=-\tau}^{\min(T - \tau, t^*)} \phi(t' + \tau, x[t]). \quad (\text{A.80})$$

CHAPTER A. APPENDICES

Proof. The index of summation is exchanged as $t' \rightarrow t - \tau$. \square

The resultant consistency constraint from Lemma A.9.2 has an identical form as (11.10):

$$\pi_{\#}^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) + \nu = S_{\#}^{\tau}(\mu_h + \pi^{tx_0}\bar{\mu}_0). \quad (\text{A.81})$$

A.9.2.4 Measure Program

The peak estimation measure LP that upper-bounds (A.76) is

$$p^* = \sup \langle p, \mu_p \rangle \quad (\text{A.82a})$$

$$\langle 1, \mu_0 \rangle = 1 \quad (\text{A.82b})$$

$$\pi_{\#}^t \mu_h = \sum_{t'=-\tau}^{-1} \delta_{t=t'} \quad (\text{A.82c})$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_{\#}^{tx_0}((t+1, f, x_1)_{\#} - Id_{\#})(\bar{\mu}_0 + \bar{\mu}_1) \quad (\text{A.82d})$$

$$\pi_{\#}^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) + \nu = S_{\#}^{\tau}(\mu_h + \pi_{\#}^{tx_0}\bar{\mu}_0) \quad (\text{A.82e})$$

$$\text{Measure Definitions from (11.3).} \quad (\text{A.82f})$$

This upper-bound also follows from constructing measures from trajectories as in Theorem 11.3.2.

A.10 Proof of Strong Duality for DDE Peak Estimation

This appendix will prove strong duality between programs (11.11) and (11.12) for peak estimation. The general pattern of Appendix A.1 will be followed.

The signed measure spaces of (11.3) are

$$\begin{aligned} \mathcal{X} &= \mathcal{M}(H_0) \times \mathcal{M}(X_0) \times \mathcal{M}([0, T] \times X)^2 \\ &\quad \times \mathcal{M}([0, T - \tau] \times X^2) \times \mathcal{M}([T - \tau, T] \times X^2) \\ \mathcal{X}' &= C(H_0) \times C(X_0) \times C([0, T] \times X)^2 \\ &\quad \times C([0, T - \tau] \times X^2) \times C([T - \tau, T] \times X^2). \end{aligned} \quad (\text{A.83})$$

CHAPTER A. APPENDICES

Their nonnegative subcones (with (11.3) membership) are topological duals under A1 and have definitions

$$\begin{aligned}\mathcal{X}_+ &= \mathcal{M}_+(H_0) \times \mathcal{M}_+(X_0) \times \mathcal{M}_+([0, T] \times X)^2 \\ &\quad \times \mathcal{M}_+([0, T - \tau] \times X^2) \times \mathcal{M}_+([T - \tau, T] \times X^2) \\ \mathcal{X}'_+ &= C_+(H_0) \times C_+(X_0) \times C_+([0, T] \times X)^2 \\ &\quad \times C_+([0, T - \tau] \times X^2) \times C_+([T - \tau, T] \times X^2).\end{aligned}\tag{A.84}$$

The collection of measures in (11.3) will be denoted as $\boldsymbol{\mu} = (\mu_h, \mu_0, \mu_p, \nu, \bar{\mu}_0, \bar{\mu}_1)$ and is a member of \mathcal{X}_+ . The constraint spaces of (11.11b)-(11.11e) are

$$\mathcal{Y} = \mathbb{R} \times C([-T, 0]) \times C^1([0, T] \times X) \times C([0, T] \times X)\tag{A.85}$$

$$\mathcal{Y}' = 0 \times \mathcal{M}([-T, 0]) \times C^1([0, T] \times X)' \times \mathcal{M}([0, T] \times X).\tag{A.86}$$

The space \mathcal{X} has the weak-* topology and \mathcal{Y} has a sup-norm bounded weak topology. Because there are no affine-inequality constraints present in (11.11b)-(11.11e), we write $\mathcal{Y}_+ = \mathcal{Y}$ and $\mathcal{Y}'_+ = \mathcal{Y}'$ to match the notation used in [40].

The variables of (11.12) with $\boldsymbol{\ell} = (\gamma, \xi, v, \phi)$ satisfy $\boldsymbol{\ell} \in \mathcal{Y}'_+$.

A pair of adjoint linear operators $\mathcal{A} : \mathcal{X}_+ \rightarrow \mathcal{Y}_+$ and $\mathcal{A}' : \mathcal{Y}'_+ \rightarrow \mathcal{X}'_+$ induced from (11.11b)-(11.11e) are

$$\begin{aligned}\mathcal{A}(\boldsymbol{\mu}) &= \begin{bmatrix} \langle 1, \mu_0 \rangle \\ \pi_\#^t \mu_h \\ \mu_p - \delta_0 \otimes \mu_0 - \mathcal{L}_f^\dagger \mu \\ S_\#^\tau (\mu_h + \pi_\#^{tx_0} \bar{\mu}_0) - \pi_\#^{tx_1} (\bar{\mu}_0 + \bar{\mu}_1) - \nu \end{bmatrix} \\ \mathcal{A}'(\boldsymbol{\ell}) &= \begin{bmatrix} \xi(t) + \phi(t + \tau, x) \\ \gamma - v(0, x) \\ v(t, x) \\ -\phi(t, x) \\ -\mathcal{L}_f v(t, x_0) - \phi(t, x_1) + \phi(t + \tau, x_0) \\ -\mathcal{L}_f v(t, x_0) - \phi(t, x_1) \end{bmatrix}.\end{aligned}\tag{A.87}$$

The cost and answer vectors are

$$\mathbf{c} = [0, 0, p, 0, 0, 0]\tag{A.88}$$

$$\mathbf{b} = [1, \lambda_{[-\tau, 0]}, 0, 0].\tag{A.89}$$

CHAPTER A. APPENDICES

Problem 11.11 may be expressed as the standard-form LP

$$p^* = \sup_{\boldsymbol{\mu} \in \mathcal{X}_+} \langle \mathbf{c}, \boldsymbol{\mu} \rangle = \langle p, \mu_p \rangle, \quad \mathbf{b} - \mathcal{A}(\boldsymbol{\mu}) \in \mathcal{Y}_+. \quad (\text{A.90})$$

The standard-form dualization of (A.90) is

$$d^* = \inf_{\boldsymbol{\ell} \in \mathcal{Y}'_+} \langle \boldsymbol{\ell}, \mathbf{b} \rangle = \gamma + \int_{t=-\tau}^0 \xi(t) dt, \quad \mathcal{A}'(\boldsymbol{\ell}) - \mathbf{c} \in \mathcal{X}_+. \quad (\text{A.91})$$

The standard-form (A.91) may be expanded into (11.12).

Given that all sets are compact (A1), measures in $\boldsymbol{\mu}$ are bounded (Lemma 11.4.2), functions in (c, \mathcal{A}) are continuous (A2, A4, $v \in C^1([0, T] \times X) \implies \mathcal{L}_f v \in C([0, T] \times X)$), and there exists a feasible measure solution (Theorem 11.3.2); it holds that strong duality between (11.11) and (11.12) is achieved by Theorem 2.6 of [40].

A.11 Subvalue Functionals and DDE Control

This appendix analyzes a DDE OCP when posed over a given history $x_h(\cdot)$. The function $J(t, x_0, x_1, u)$ is a running cost evaluated on the trajectory starting from x_h , and $J_T(x)$ is a terminal cost at time T . The final point $x(T)$ must reside in the terminal set $X_T \subseteq X$. The controller $u(\cdot)$ must reside inside the compact set $U \subset \mathbb{R}^m$ at each time. The DDE OCP under these constraints is

$$P^* = \inf_{u(t)} \int_{t=0}^T J(t, x(t), x(t-\tau), u(t)) dt + J_T(x(T)) \quad (\text{A.92a})$$

$$\dot{x} = f(t, x(t), x(t-\tau), u(t)) \quad \forall t \in [0, T] \quad (\text{A.92b})$$

$$u(t) \in U \quad \forall t \in [0, T] \quad (\text{A.92c})$$

$$x(t) = x_h(t) \quad \forall t \in [-\tau, 0] \quad (\text{A.92d})$$

$$x(T) \in X_T. \quad (\text{A.92e})$$

Problem A.92 was addressed in works such as [159, 161, 163, 162, 195], and was completely solved in the case of linear DDE dynamics and a quadratic objective in [196, 197].

A pair of infinite-dimensional LPs are synthesized to bound the OCP in (A.92).

This appendix assumes that the terminal time T is fixed to simplify analysis. The MV-solution from Section 11.3 involves a free terminal time, multiple histories, and zero running cost ($J = 0$).

CHAPTER A. APPENDICES

A.11.1 Control Measure Program

The deterministic control law $u(t, x_0, x_1)$ at each time t is relaxed into a probability distribution $\xi_u(u \mid t, x_0, x_1)$ [45].

The measures involved in an MV-solution of (A.92) are

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (\text{A.93a})$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.93b})$$

$$\text{Occupation Start} \quad \bar{\mu}_0 \in \mathcal{M}_+([0, \kappa T] \times X^2 \times U) \quad (\text{A.93c})$$

$$\text{Occupation End} \quad \bar{\mu}_1 \in \mathcal{M}_+([\kappa T, T] \times X^2 \times U). \quad (\text{A.93d})$$

The time-slack measure is set to $\nu = 0$ because of the fixed-terminal-time setting ($t^* = T$). The symbol $\mu_{x_h(\cdot)}$ is the occupation measure of $t \mapsto (t, x_h(t))$ between $t \in [-\tau, 0]$. The history measure μ_h from (11.3) is set equal to $\mu_{x_h(\cdot)}$ in the case of a single history. Similarly, the initial measure μ_0 is set to the Dirac delta $\delta_{x=x_h(0^+)}$.

A measure relaxation to the optimal program in (A.92) is

$$p^* = \inf \langle J, \bar{\mu} \rangle + \langle J_T, \mu_T \rangle \quad (\text{A.94a})$$

$$\delta_T \otimes \mu_T = \delta_{t=0, x=x_h(0^+)} + \pi_\#^{tx_0} \mathcal{L}_f^\dagger(\bar{\mu}_0 + \bar{\mu}_1) \quad (\text{A.94b})$$

$$\pi_\#^{tx_1}(\bar{\mu}_0 + \bar{\mu}_1) = S_\#^\tau(\mu_{x_h(\cdot)} + \pi_\#^{tx_0} \bar{\mu}_0) \quad (\text{A.94c})$$

$$\text{Measures from (A.93).} \quad (\text{A.94d})$$

This measure relaxation is based on the optimal control framework of [7, 8]. Young measure formulations for DDE OCP have been developed in [159, 161, 163, 162, 160], but Liouville equations began use only in [170].

A.11.2 Control Function Program

Dual variables $v \in C^1([0, T] \times X)$ and $\phi \in C([0, T] \times X)$ may be introduced to form the dual program of (A.94):

$$d^* = \sup v(0, x_h(0)) + \int_{-\tau}^0 \phi(t + \tau, x_h(t)) dt \quad (\text{A.95a})$$

$$J_T(x) - v(T, x) \geq 0 \quad \forall x \in X_T \quad (\text{A.95b})$$

$$\mathcal{L}_f v + J(t, x_0, u) - \phi(t, x_1) + \phi(t + \tau, x_0) \geq 0 \quad \forall (t, x_0, x_1, u) \in [0, T - \tau] \times X^2 \times U \quad (\text{A.95c})$$

$$\mathcal{L}_f v + J(t, x_0, u) - \phi(t, x_1) \geq 0 \quad \forall (t, x_0, x_1, u) \in [T - \tau, T] \times X^2 \times U \quad (\text{A.95d})$$

$$v \in C^1([0, T] \times X) \quad (\text{A.95e})$$

$$\phi \in C([0, T] \times X). \quad (\text{A.95f})$$

This dual program is obtained (with strong duality) by following nearly identical steps to Appendix A.10.

A.11.3 True Value Functional

Let \mathcal{U} be the admissible class of control inputs (such as $\mathcal{U} = \{u : [0, T] \rightarrow U\}$). Given a time $t \in [0, T]$, a current state $z \in X$, and a history $w(\cdot) \in PC([- \tau, 0], X)$, the value functional V^* associated with the OCP (A.92) is

$$\begin{aligned} V^*(t, z, w(\cdot)) &= \min_{u \in \mathcal{U}} \int_{t' = t}^T J(t', x(t' | t, w, u), u(t')) dt' + J_T(x(T | t, w, u)) \\ &\quad \dot{x} = f(t, x(t), x(t - \tau), u(t)) \quad \forall t \in [0, T] \\ &\quad x(t') = w(t') \quad \forall t \in [t - \tau, t) \\ &\quad x(t) = z \\ &\quad x(T) \in X_T \\ &\quad u(t) \in U \quad \forall t \in [0, T]. \end{aligned} \quad (\text{A.96})$$

The value functional V^* is the cost of solving Problem (A.92) starting at time t and state z with history $w(\cdot)$. The convention of arguments $t, z, w(\cdot)$ was taken from [195].

CHAPTER A. APPENDICES

The Hamilton-Jacobi-Bellman (HJB) equation of optimality is [107]

$$0 = J_T(x(T)) - V^*(T, x(T)) \quad (\text{A.97a})$$

$$0 = \inf_{u \in U} \left(\dot{V}^*(t, x(t), x_\tau(\cdot), u) + J(t, x(t), u(t)) \right) \quad \forall t \in [0, T]. \quad (\text{A.97b})$$

The Cauchy problem of (7, 8) in [195] has the form of (A.97).

A.11.4 Subvalue Functional

The solution of program (A.95) can create a lower-bound on the value functional V^* .

A.11.4.1 Properties of Subvalue Functionals

Definition A.11.1. A subvalue functional is a functional $\mathcal{V}(t, x, w)$ such that

$$\mathcal{V}(t, x, w) \leq V^*(t, x, w) \quad \forall t \in [0, T], x \in X, w \in PC([-\tau, 0], X). \quad (\text{A.98})$$

Theorem A.11.1. Any functional $\mathcal{V}(t, x, x_\tau)$ with derivative $\dot{\mathcal{V}}(t, x, x_\tau, u)$ that satisfies the following two properties is a subvalue functional for V^* :

$$J_T(x) - \mathcal{V}(T, x, w) \geq 0 \quad \forall x \in X_T, w \in PC([-\tau, 0], X) \quad (\text{A.99a})$$

$$J(t, x, u) + \dot{\mathcal{V}}(t, x, w, u) \geq 0, \quad \forall t \in T, x \in X, w \in PC([-\tau, 0], X), u \in U. \quad (\text{A.99b})$$

Proof. This result follows by following the steps of Proposition 1's proof from [106].

Let $\tilde{u} \in \mathcal{U}$ be an arbitrary control policy starting at the initial condition (t_0, z, w) , resulting in a trajectory $\tilde{x}(t)$. Denote $\tilde{x}_\tau(\cdot)$ as the history function $\tilde{x}_t(t') = \tilde{x}(t + t') \forall t' \in [-\tau, 0]$.

Relation (A.99b) ensures that for all $t \in [t_0, T]$,

$$\dot{\mathcal{V}}(t, \tilde{x}(t), \tilde{x}_t(\cdot), \tilde{u}(t)) + J(t, \tilde{x}(t), \tilde{u}(t)) \geq 0. \quad (\text{A.100})$$

Integrating the above term with respect to t yields

$$0 \leq \int_{t=t_0}^T \dot{\mathcal{V}}(t, \tilde{x}(t), \tilde{x}_t(\cdot), \tilde{u}(t)) + J(t, \tilde{x}(t), \tilde{u}(t)) dt \quad (\text{A.101a})$$

$$0 \leq \underbrace{\mathcal{V}(T, \tilde{x}(T), \tilde{x}_T(\cdot))}_{V(T)} - \underbrace{\mathcal{V}(t_0, \tilde{x}(t_0), \tilde{x}_{t_0}(\cdot))}_{\mathcal{V}(t_0)} + \int_{t=t_0}^T J(t, \tilde{x}(t), \tilde{u}(t)) dt \quad (\text{A.101b})$$

$$\mathcal{V}(t_0) \leq V(T) + \int_{t=t_0}^T J(t, \tilde{x}(t), \tilde{u}(t)) dt \quad (\text{A.101c})$$

$$\mathcal{V}(t_0) \leq J_T(\tilde{x}(T)) + \int_{t=t_0}^T J(t, \tilde{x}(t), \tilde{u}(t)) dt. \quad (\text{A.101d})$$

CHAPTER A. APPENDICES

The transformation of (A.101c) to (A.101d) follows from relations (A.103) and (A.95b) ($\mathcal{V}(T) \leq J(\tilde{x}(T))$). When \tilde{u} is a minimizing control u^* (if it exists), then the right-hand side of (A.101d) is the optimal value functional $V^*(t_0, z, w(\cdot))$ and the left-hand side is $\mathcal{V}(t_0, z, w(\cdot))$. The proof that \mathcal{V} is a lower bound on V^* is therefore complete, given that $\mathcal{V}(t_0, z, w(\cdot)) \leq V^*(t_0, z, w(\cdot))$ will hold for all choices of $(t_0, z, w(\cdot))$. \square

A.11.4.2 Recovery of a Subvalue Functional

The dual solution (v, ϕ) from (A.95) may be assembled into a functional,

$$V(t, x, z(\cdot)) = v(t, x) + \int_t^{\min(t+\tau, T)} \phi(s, z(s-\tau)) ds. \quad (\text{A.102})$$

The bias term ϕ in (A.95f) is defined and is C^0 continuous only between times $t \in [0, T]$. If the hard integration limit at T was not present, then ϕ would be queried at undefined values $t \in (T, T + \tau]$. The terminal value of the value functional is

$$V(T) = V(t, x(T), x([T-\tau, T])) = v(T, x(T)) + \int_T^T \phi_i(s, x(s-\tau)) ds = v(T, x(T)). \quad (\text{A.103})$$

The objective in (A.95a) is the evaluation of the value functional at time $t = 0$ along the optimal controlled trajectory $x^*(t)$:

$$V(0) = V(0, x_h(0), x_h) = v(0, x_h(0)) + \int_0^\tau \phi_i(s, x_h(s-\tau)) ds. \quad (\text{A.104})$$

The time-derivative (co-invariant derivative) of the value functional \dot{V} is

$$\dot{V}(t, z, w(\cdot), u) = \mathcal{L}_f v(t, x(t)) + I_{[0, T-\tau]}(t) \phi(t + \tau, z) - \phi(t, w(-\tau)). \quad (\text{A.105})$$

Theorem A.11.2. *The functional (A.102) is a subvalue functional in the sense of (A.11.1).*

Proof. The terminal constraint (A.95b) satisfies (A.99a) given the terminal value evaluation in (A.103). The combination of (A.95c) and (A.95d) together satisfy (A.99b) under the derivative value expression in (A.105). \square

A.11.4.3 Continuity of the Recovered Value Functional

Let $u^*(t) \in \mathcal{U}$ be the optimal (infimizing) trajectory of problem (A.92) given $z, w(\cdot)$, inducing a controlled trajectory $x^*(t) = x(t \mid z, w(\cdot))$. The value functional evaluated along the

CHAPTER A. APPENDICES

optimal trajectory $V^*(t) = V(t, x^*(t), x^*([t - \tau, t]))$ is C^0 -continuous in t , but is not necessarily C^1 -continuous in t . At the time $t = T - \tau$, define the value functional evaluations

$$V_-^* = \lim_{t \rightarrow (T-\tau)^-} V^*(t) \quad V_+^* = \lim_{t \rightarrow (T-\tau)^+} V^*(t). \quad (\text{A.106})$$

The difference in these evaluations for every lag i^* is

$$\Delta V^* = V_-^* - V_+^* = \left(\int_{T-\tau^-}^{T^-} - \int_{T-\tau^+}^{T^+} \right) \phi(s, x(s - \tau_i)) ds = 0. \quad (\text{A.107})$$

The value functional $V^*(t)$ is therefore a member of $C^0([0, T])$ given that $v \in C^1([0, T] \times X)$ and $\phi \in C^0([0, T] \times X)$.

The value functional derivative evaluations are

$$\dot{V}_-^* = \lim_{t \rightarrow (T-\tau)^-} \dot{V}^*(t) \quad \dot{V}_+^* = \lim_{t \rightarrow (T-\tau)^+} \dot{V}^*(t). \quad (\text{A.108})$$

The difference in the derivative evaluations on both sides of $t = T - \tau$ is

$$\Delta \dot{V}^* = \dot{V}_-^* - \dot{V}_+^* = \phi(T, x(T - \tau)). \quad (\text{A.109})$$

It is not guaranteed that $\phi(T, x(T - \tau)) = 0$, so the value functional V^* may have discontinuous first derivatives. The value functional is therefore C^0 in time along trajectories, and fails to be C^1 at the time $t = T - \tau$.

We form an additional conjecture to 11.3.1 in the OCP case based on the tightness conditions in [7].

Conjecture A.11.1. *Assume for the purposes of this conjecture that:*

A1' *The sets $\{[-\tau, T], X, X_T, U\}$ are all compact.*

A2' *The costs J, J_T are continuous.*

A3' *The dynamics $f(t, x_0, x_1, u)$ are Lipschitz in their arguments.*

A4' *The history x_h is inside $PC([-\tau, 0], X)$.*

A5' *The image of $f(t, x_0, X, U)$ is convex for each fixed (t, x_0) .*

A6' *The mapping $v \mapsto \inf_{u \in U} J(t, x_0, x_1, u) : f(t, x_0, x_1, u) = v$ is convex in $v \in \mathbb{R}^n$.*

Then there is no relaxation gap between (A.92) and (A.95) ($P^ = p^*$).*

A.11.5 Approximate Recovery

A control policy $u(t)$ may be extracted from the value functional V through the trajectory condition (A.97b) by

$$u(t) = \underset{u}{\operatorname{argmin}} \mathcal{L}_f v(t, x(t)) + I_{[0, T-\tau]}(t) \phi(t + \tau, x(t)) * -\phi(t, x(t - \tau)) + J(t, x(t), u(t)) \quad (\text{A.110a})$$

$$= \underset{u}{\operatorname{argmin}} f(t, x(t), x(t - \tau), u) \cdot \nabla_x v(t, x(t)) + J(t, x(t), u). \quad (\text{A.110b})$$

The work in [106] quantifies performance bounds of polynomial value function approximations for ODE systems in terms of W^1 Sobolev norms away from the true value function. Quantifying performance bounds of the extracted controller of (A.110a) is an open problem.

A.11.6 Example of Optimal Control

An example of optimal control is presented on the one-dimensional linear system:

$$\begin{aligned} x'(t) &= -3x(t) - 5x(t - 0.25) + u & \forall t \in [0, 1] \\ x_h(t) &= -1 & \forall t \in [-0.25, 0]. \end{aligned} \quad (\text{A.111})$$

This system has one lag with $\tau = 0.25$ and a time horizon of $T = 1$. The state and control constraints are $X = [-1, 1]$ and $U = [-1, 1]$. With a control weight of $R = 0.01$, the penalties are

$$J(t, x, u) = 0.5x^2 + 0.5Ru^2 \quad J_T(x) = 0. \quad (\text{A.112})$$

The open loop total cost is 0.0674. Table A.1 lists optimal control value approximations for this system.

Table A.1: SDP approximation bounds to program (A.94)

order	1	2	3	4	5	6
bound	7.90E-05	0.0322	0.0386	0.0391	0.0393	0.0393

The applied control $u(t)$ may be recovered through equation (A.110a) as

$$u(t) = \operatorname{Saturate}_{[-1, 1]} \left(-\frac{1}{R} \partial_x v(t, x(t)) \right). \quad (\text{A.113})$$

The trajectories and nonnegative functions are plotted for order 4 in Figures A.1-A.2. The order 4 control bound is 0.0391, and the cost evaluated along the controlled trajectory is 0.0394.

CHAPTER A. APPENDICES

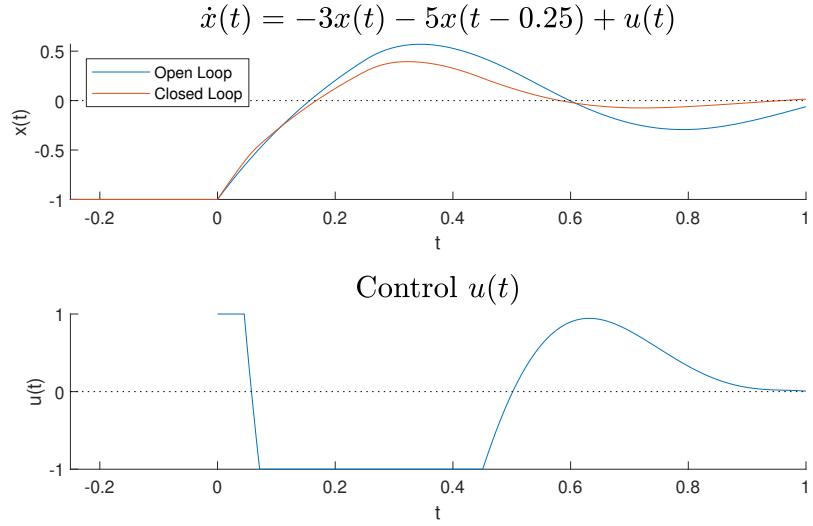


Figure A.1: Open and closed-loop trajectory with control

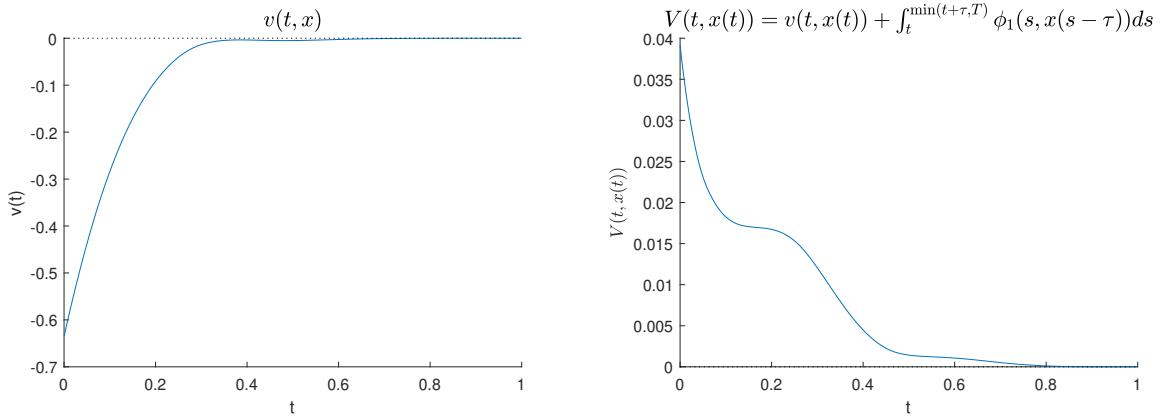


Figure A.2: Auxiliary function v and value functional V from (A.102)

A.12 Improved Accuracy of DDE Control Problems

This appendix lays out methods to reduce the conservatism of DDE OCPs from appendix A.11 by adding new infinite-dimensional nonnegativity constraints.

All approaches discussed in this appendix may be applied to peak estimation, but are presented here in the simplified fixed-terminal-time single-history OCP setting.

A.12.1 Spatial Partitioning

The constraints (A.95b)-(A.95d) must hold in support sets defined by $[0, T] \times X \times U$. Assume that there exists a decomposition of the state spaces $X = \cup_k X_k$ such that $\forall k : \dim(X_k) = n$ and $\forall k, k' : \text{int}(X_k \cap X_{k'}) = \emptyset$ (cells X_k are full-dimensional and their intersections are not full dimensional). Further assume a similar decomposition exists for the control set $U = \cup_\ell U_\ell$.

Let $(v_k(t, x), \phi_k(t, x))$ be functions associated with each space X^k . A space-control partition of (A.95) is:

$$d^* = \sup \sum_k (I_{X_k}(x_h(0))v_k(0, x_h(0))) + \int_{-\tau}^0 \phi(t + \tau, x_h(t))dt \quad (\text{A.114a})$$

$$\forall x \in X_k :$$

$$J_T(x) - v_k(T, x) \geq 0 \quad (\text{A.114b})$$

$$\forall (t, x_0, x_1, u) \in [0, T - \tau] X_k \times X_{k'} \times U_\ell :$$

$$\mathcal{L}_f v_k + J(t, x_0, u) - \phi_{k'}(t, x_1) + \phi_k(t + \tau, x_0) \geq 0 \quad (\text{A.114c})$$

$$\forall (t, x_0, x_1, u) \in [T - \tau, T] \times X_k \times X_{k'} \times U_\ell :$$

$$\mathcal{L}_f v_k + J(t, x_0, u) - \phi_{k'}(t, x_1) \geq 0 \quad (\text{A.114d})$$

$$\forall (t, x) \in [0, T] \times (X_k \cap X_{k'}) :$$

$$v_k(t, x) = v_{k'}(t, x) \quad (\text{A.114e})$$

$$\forall k : v_k \in C^1([0, T] \times X_k) \quad (\text{A.114f})$$

$$\forall k : \phi_k \in C([0, T] \times X_k). \quad (\text{A.114g})$$

The v_k terms agree on boundary regions between state cells by (A.114e). The ϕ_k terms remain continuous (bounded measurable), but this partitioning has an impact when evaluating the finite-degree SDPs.

A.12.2 Temporal Partitioning

We utilize the following lemma to provide conditions for temporal partitioning.

Lemma A.12.1. *A sufficient condition for $\int_{t_0}^{t_1} g_1(t)dt \geq \int_{t_0}^{t_1} g_2(t)dt$ is that*

$$\forall t \in [t_0, t_1] : g_1(t) \geq g_2(t). \quad (\text{A.115})$$

Define the following time breaks (partition) arranged in sorted order as

$$T_{break} = \{0, t_1, \dots, t_{k-1}, t_k = T - \tau, t_{k+1}, \dots, t_{k+\ell-1}, t_{k+\ell} = T\}. \quad (\text{A.116})$$

CHAPTER A. APPENDICES

Let $[t_b, t_{b+1}]$ and $[t_{b-1}, t_b]$ be regions in T_{break} . The subvalue functionals from (A.102) defined in this region must satisfy

$$V_b(t_b, x, z(\cdot)) \geq V_{b-1}(t_b, x, z(\cdot)) \quad (\text{A.117a})$$

$$v_b(t_b, x) + \int_{t_b}^{\min(t_b+\tau, T)} \phi_b(s, z(s-\tau)) ds \geq v_{b-1}(t_b, x) + \int_{t_b}^{\min(t_b+\tau, T)} \phi_{b-1}(s, z(s-\tau)) ds. \quad (\text{A.117b})$$

The sufficient condition in Lemma A.12.1 may be used to accomplish this relation in (A.117), ensuring that the subvalue function will always decrease when traversing a time break:

$$v_b(t_b, x) \geq v_{b-1}(t_b, x) \quad \forall x \in X \quad (\text{A.118a})$$

$$\phi_b(t, x) \geq \phi_{b-1}(t, x) \quad \forall (t, x) \in [t_b, \min(t_b + \tau, T)] \times X. \quad (\text{A.118b})$$

The resultant time-partitioned LP is

$$d^* = \sup \quad v(0, x_h(0)) + \int_{-\tau}^0 \phi_{k+\ell}(t + \tau_i, x_h(t)) dt \quad (\text{A.119a})$$

$$\forall x \in X :$$

$$J_T(x) - v_{k+\ell}(T, x) \geq 0 \quad (\text{A.119b})$$

$$\forall (t, x_0, x_1, u) \in [t_{k'}, t_{k'+1}] \times X^2 \times U, k' = 0..k-1 :$$

$$\mathcal{L}_f v_{k'} + J(t, x_0, u) - \phi_{k'}(t, x_1) + \phi_{k'}(t + \tau, x_0) \geq 0 \quad (\text{A.119c})$$

$$\forall (t, x_0, x_1, u) \in [t_{k'}, t_{k'+1}] \times X^2 \times U, k' = k..k+\ell :$$

$$\mathcal{L}_f v_{k'} + J(t, x_0, u) - \phi_{k'}(t, x_1) \geq 0 \quad (\text{A.119d})$$

$$\forall x \in X, k' = 1..k+\ell-1 :$$

$$v_{k'}(t_{k'}, x) \leq v_{k'+1}(t_{k'}, x) \quad (\text{A.119e})$$

$$\forall (t, x) \in [t_{k'}, \min(t_{k'} + \tau, T)] \times X, k' = 1..k+\ell-1$$

$$\phi_{k'}(t, x) \leq \phi_{k'+1}(t, x) \quad (\text{A.119f})$$

$$\forall k' = 0..k+\ell : \quad (\text{A.119g})$$

$$v_{k'} \in C^1([t_{k'}, t_{k'+1}] \times X) \quad (\text{A.119h})$$

$$\phi_{k'} \in C([t_{k'}, \min(t_{k'+1} + \tau, T)] \times X). \quad (\text{A.119i})$$

A.12.3 Double Integral Functionals

The subvalue functional in (A.102) has a single integral term for each delay. Some Lyapunov-Krasovskii or Barrier methods for DDE analysis employ double integrals, such as the

CHAPTER A. APPENDICES

following functional for a single delay τ [164]:

$$V(t, z, w) = v(t, z) + \int_t^{\min(t+\tau, T)} \phi_i(s, w(s-\tau-t)) ds \\ + \int_t^{\min(t+\tau, T)} \int_{-\tau}^0 \psi(s, q, w(s-\tau-t), w(q)) dq ds. \quad (\text{A.120})$$

The time derivative of (A.120) in the time span $t \in [0, T - \tau]$ is

$$\dot{V}(t, z, w) = \mathcal{L}_f v(t, z) + \phi(t + \tau, w(0)) - \phi(t, w(-\tau)) \quad (\text{A.121})$$

$$+ \int_{-\tau}^0 \psi(t + \tau, q, w(0), w(q)) dq \\ - \int_{-\tau}^0 \psi(t, q, w(-\tau), w(q)) dq, \quad (\text{A.122})$$

and between $t \in (T - \tau, 0]$ is

$$\dot{V}(t, z, w) = \mathcal{L}_f v(t, z) - \phi(t, w(-\tau)) - \int_{-\tau}^0 \psi(t, q, w(-\tau), w(q)) dq. \quad (\text{A.123})$$

The derivative \dot{V} has a discontinuity present at $t = T - \tau$, just as described in Section (A.11.4.3).

A sufficient condition for the inequality (A.99b) to be fulfilled is that the following functions associated with (A.122) and (A.123) (moving all terms under the dq integral) are nonnegative:

$$\forall t \in [0, T - \tau], (x_0, x_1, \tilde{x}) \in X^3, q \in [-\tau, 0] : \\ \tau^{-1} (\mathcal{L}_f v(t, z) + J(t, x_0, u) + \phi(t + \tau, x_0) - \phi(t, x_1)) \\ + \psi(t + \tau, q, x_0, \tilde{x}) - \psi(t, q, x_0, \tilde{x}) \geq 0 \quad (\text{A.124})$$

$$\forall t \in [T - \tau, T], (x_0, x_1, \tilde{x}) \in X^3, q : \\ \tau^{-1} (\mathcal{L}_f v(t, z) + J(t, x_0, u) - \phi(t, x_1)) - \psi(t, q, x_0, \tilde{x}) \geq 0. \quad (\text{A.125})$$

Lemma A.12.1 is utilized to enforce nonnegativity of the integral terms in (A.122) and (A.123). The τ^{-1} scale factor arises from placing a q -independent term (such as $J(t, x_0, u)$) inside the integral. The variable $q \in [-\tau, 0]$ is the integrated (swept) time, and $\tilde{x} \in X$ abstracts out the swept state $w(q)$. The dual formulation of constraints (A.124) and (A.125) involve occupation measures $\bar{\mu}_0 \in \mathcal{M}_+([0, T - \tau] \times X^3 \times U)$ and $\bar{\mu}_1 \in \mathcal{M}_+([T - \tau, T] \times X^3 \times U)$. This construction may be generalized to DDEs with r delays by adding a double-integral term for each delay.

A.13 Joint + Component Measure

This appendix details an alternate notion of MV-solutions for DDEs. Solving peak estimation problems through these methods will return more conservative but quicker-executing programs (computationally) as compared to the results in Section 11.3.

A.13.1 Measure Program

The MV-solution involves a joint occupation measure $\bar{\mu}$ and component measures ω_0, ω_1 :

$$\text{History} \quad \mu_h \in \mathcal{M}_+(H_0) \quad (\text{A.126a})$$

$$\text{Initial} \quad \mu_0 \in \mathcal{M}_+(X_0) \quad (\text{A.126b})$$

$$\text{Peak} \quad \mu_p \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.126c})$$

$$\text{Time-Slack} \quad \nu \in \mathcal{M}_+([0, T] \times X) \quad (\text{A.126d})$$

$$\text{Joint Occupation} \quad \bar{\mu} \in \mathcal{M}_+([0, T] \times X^2) \quad (\text{A.126e})$$

$$\text{Component Start} \quad \omega_0 \in \mathcal{M}_+([0, T - \tau] \times X) \quad (\text{A.126f})$$

$$\text{Component End} \quad \omega_1 \in \mathcal{M}_+([T - \tau, T] \times X). \quad (\text{A.126g})$$

The peak estimation LP for the Joint+Component framework is

$$p^* = \sup \langle p, \mu_p \rangle \quad (\text{A.127a})$$

$$\langle 1, \mu_0 \rangle = 1 \quad (\text{A.127b})$$

$$\pi_\#^t \mu_h = \lambda_{[-\tau, 0]} \quad (\text{A.127c})$$

$$\mu_p = \delta_0 \otimes \mu_0 + \pi_\#^{tx_0} \mathcal{L}_f^\dagger \bar{\mu} \quad (\text{A.127d})$$

$$\pi_\#^{tx_0} \bar{\mu} = \omega_0 + \omega_1 \quad (\text{A.127e})$$

$$\pi_\#^{tx_1} \bar{\mu} + \nu = S_\#^\tau(\mu_h + \omega_0) \quad (\text{A.127f})$$

$$\text{Measure Definitions from (A.126).} \quad (\text{A.127g})$$

The history-validity and Liouville constraints in (A.127) are the same as in (11.11) under the relation $\bar{\mu} = \bar{\mu}_0 + \bar{\mu}_1$. The consistency constraint in the Joint+Component formulation is split up into the pair (A.127e)-(A.127f).

Theorem A.13.1. *Program (A.127) returns an upper bound on (11.1).*

CHAPTER A. APPENDICES

Proof. Let $x_h \in \mathcal{H}$ be a history that generates the trajectory $x(t | x_h)$, and let $t^* \in [0, T]$ be a stopping time.

Just as in Theorem 11.3.2, measures can be picked as $\mu_0 = \delta_{x=x_h(0^+|x_h)}$, $\mu_p = \delta_{t=t^*} \otimes \delta_{x=x(t^*|x_h)}$, μ_h as the occupation measure of $x_\xi(t)$ in the times $[-\tau, 0]$, and $\bar{\mu}$ as the occupation measure of $z(t) = (x(t | x_h), x(t - \tau | x_h))$ in the times $[0, t^*]$.

When $t^* \in [0, T - \tau]$, then ω_0 is the occupation measure of $x(t | x_h)$ in times $[0, t^*]$, ω_1 is the zero measure, and ν is the occupation measure of $x(t - \tau | x_h)$ in times $[t^*, t^* + \tau]$. When $t^* \in (T - \tau, T]$, then ω_0 is the occupation measure of $x(t | x_h)$ in the times $[0, T - \tau]$, ω_1 is the occupation measure of $x(t | x_h)$ in the times $[T - \tau, t^*$, and ν is the occupation measure of $x(t - \tau | x_h)$ in the times $[T - \tau, T]$. All measures inside (A.126) have been defined for a valid trajectory, proving that (A.127) upper-bounds (11.1). \square

Theorem A.13.2. *The Joint+Component (A.127) is also an upper bound on (11.11).*

Proof. Let $(\mu_h, \mu_0, \mu_p, \nu, \bar{\mu}_0, \bar{\mu}_1)$ be a feasible set of measures for the constraints of (11.11).

After performing the following definitions,

$$\bar{\mu} = \bar{\mu}_0 + \bar{\mu}_1 \quad \omega_0 = \pi^{tx_0} \bar{\mu}_0 \quad \omega_1 = \pi^{tx_0} \bar{\mu}_1, \quad (\text{A.128})$$

the measures $(\mu_h, \mu_0, \mu_p, \nu, \bar{\mu}, \omega_0, \omega_1)$ are feasible solutions for the constraints of (A.127). \square

Note how the Joint+Component MV-solution involves only one measure involving (t, x_0, x_1) together ($\bar{\mu}$ in (A.126d)), while the solution in (11.3) has two measures $(\bar{\mu}_0, \bar{\mu}_1)$. Application of the moment-SOS hierarchy towards solving problems in (A.126d) result in only one Gram matrix of maximal size $\binom{1+2n+d}{d}$.

CHAPTER A. APPENDICES

A.13.2 Function Program

The gap between (A.127) and (11.11) can most easily be observed by examining the dual program of (A.127):

$$d^* = \inf_{\gamma \in \mathbb{R}} \gamma + \int_{t=-\tau}^0 \xi(t) dt \quad (\text{A.129a})$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (\text{A.129b})$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (\text{A.129c})$$

$$\xi(t) + \phi_1(t + \tau, x) \quad \forall (t, x) \in H_0 \quad (\text{A.129d})$$

$$0 \geq \mathcal{L}_f v(t, x_0) + \phi_0(t, x_0) + \phi_1(t, x_1) \quad \forall (t, x_0, x_1) \in [0, T] \times X^2 \quad (\text{A.129e})$$

$$\phi_1(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \quad (\text{A.129f})$$

$$\phi_0(t, x) + \phi_1(t + \tau, x) \geq 0 \quad \forall (t, x) \in [0, T - \tau] \times X \quad (\text{A.129g})$$

$$\phi_0(t, x) \geq 0 \quad \forall (t, x) \in [T - \tau, T] \times X \quad (\text{A.129h})$$

$$v(t, x) \in C^1([0, T] \times X) \quad (\text{A.129i})$$

$$\phi_0(t, x), \phi_1(t, x) \in C([0, T] \times X) \quad (\text{A.129j})$$

$$\xi(t) \in C([-\tau, 0]). \quad (\text{A.129k})$$

Adding together (A.129e) and (A.129g) yields constraint (11.12f) in $[0, T - \tau] \times X^2$.

Similarly, the addition of (A.129e) and (A.129h) forms constraint (11.12g). The dual formulation in (A.129) enforces nonnegativity of addends inside whole-terms of (11.11). The constraints of (A.129) are stricter than of (11.12), resulting in a lowered infimum/upper bound on peak value.

A.13.3 Delayed Flow System Example

Table (A.2) compares moment-SOS SDPs associated to programs (11.11) and (A.127) for applying peak estimation to the delayed Flow example in Section 11.6.2.

Table A.2: Comparison of (11.11) and (A.127) SDP bounds for the delayed Flow system

degree d	1	2	3	4	5
Joint+Component (A.127)	1.25	1.223	1.1937	1.1751	1.1636
Standard (11.11)	1.25	1.2183	1.1913	1.1727	1.1630

CHAPTER A. APPENDICES

Table A.3: Time (seconds) to obtain SDP bounds in Table A.2

degree d	1	2	3	4	5
Joint+Component (A.127)	0.782	0.991	5.271	31.885	336.509
Standard (11.11)	0.937	1.190	9.508	105.777	552.496